



Universidade do Porto

FEP Faculdade de
Economia

TESE DE MESTRADO EM ANÁLISE DE DADOS E SISTEMAS DE APOIO À DECISÃO

PREVISÃO DA ENERGIA EÓLICA - SANTIAGO - CABO VERDE

Autor

Luís Monteiro Costa

Orientador

Professor Doutor João Gama

Coorientador

Engenheiro João Sousa

13-09-2013

Nota Biográfica

Luís Monteiro Costa nasceu em Cabo Verde, na então Vila do Tarrafal de Santiago, atual cidade do Tarrafal de Santiago.

Bacharel em Matemática, no Curso de Formação de Professores para o Ensino Secundário, pelo Instituto Superior de Educação, na Cidade da Praia, Cabo Verde entre 1992 e 1995. Licenciou-se em Tecnologias de Construção Civil, área de Ensino, pelo Centro Federal de Educação Tecnológica, Rio de Janeiro, entre 1999 e 2001. Fez Pós-Graduação em Administração Escolar, pela Universidade Cândido Mendes do Rio de Janeiro em 1999.

Iniciou a sua atividade profissional no ano letivo 1992/93 como professor do Ensino Secundário na Escola Secundária do Tarrafal de Santiago, de outubro de 1992 até janeiro de 1993 onde foi transferido para a Escola Preparatória de São Miguel, ano letivo 1992/93.

Exerceu funções docentes na Escola Secundária Polivalente Cesaltina Ramos, (Escola Técnica da Praia), onde lecionou a disciplina de Matemática, do ano letivo 1995/96 a 1998/99.

Desempenhou funções de coorientador de estágio pedagógico na área de Matemática dos formandos finalistas do Instituto Superior de Educação, no ano letivo 1996/97 e de coordenador pedagógico da disciplina de Matemática na Escola Técnica da Praia.

Nos anos letivos 2001/02 a 2003/04 lecionou a disciplina de Tecnologias de Construção Civil na Escola Técnica “Grão Duque Henri”, Assomada onde também exerceu funções de coordenador pedagógico da disciplina de Tecnologias de Construção Civil.

Coordenou o projeto de revisão do Plano Curricular do Ensino Técnico na área de Construção Civil em 2004. No mesmo ano dirigiu a equipa de elaboração dos programas do curso de formação de monitores, na referida área, para as Escolas Técnicas de São Vicente e Santa Catarina de Santiago.

Exerceu funções de Diretor da Escola Secundária do Tarrafal de Santiago de setembro de 2004 a julho de 2011. Paralelamente exerceu funções de Deputado Municipal pelo círculo eleitoral do Tarrafal de Santiago entre 2008 a 2011.

Agradecimentos

A elaboração desta dissertação não teria sido possível sem a colaboração e apoio de algumas pessoas, a quem gostaria de agradecer. Em primeiro lugar, ao Professor Doutor João Gama e ao Engenheiro João Sousa, pela orientação na elaboração deste trabalho e pela disponibilidade sempre demonstrada. A eles agradeço também a revisão muito cuidada de todos os documentos elaborados no percurso desta tese e os comentários críticos e construtivos que permitiram enriquecer o seu conteúdo.

Segundo, à empresa cabo-verdiana Cabeólica e o seu colaborador Valdemar Lopes que disponibilizou os dados analisados neste trabalho. Uma nota de apreço à Faculdade de Economia da Universidade do Porto, onde decorreu toda a minha formação académica por todos os meios colocados à minha disposição. Dirijo uma palavra de agradecimento aos meus colegas de formação pelos apoios que me têm concedido. Não posso deixar de mencionar a minha família, pela motivação, apoio constante e compreensão. Em particular, tenho que agradecer aos meus pais, às minhas irmãs e aos meus filhos.

Resumo

Neste trabalho procura-se utilizar modelos de previsão de séries temporais para prever a produção da energia elétrica a partir da energia eólica em Cabo Verde, particularmente na ilha de Santiago. É um problema que tem recebido especial atenção dos pesquisadores nos últimos anos. Prever o futuro, e em especial o comportamento de séries temporais, é fundamental em análises e apoio à tomada de decisões, e continua sendo um desafio para a estatística e para computação.

Foram utilizados modelos, *Holt-Winters*, *ARIMA* e redes neuronais artificiais, Função de Base Radial (*RNAs-RBF*) e Perceptron de múltiplas camadas (*RNAs-MLP*). O modelo *Holt-Winters* é um modelo de previsão exponencial, conhecido por lidar com elementos de tendência e sazonalidade de uma série temporal. O modelo *ARIMA* que possui apenas uma variável, descreve o comportamento de uma variável em termos de seus valores passados. As redes neurais têm-se mostrado grandes ferramentas na aplicação de previsões de séries temporais. Neste contexto, neste trabalho propõe-se a realização de uma análise comparativa desses modelos não-lineares para a previsão, tentando encontrar qual o modelo que melhor se adapta à série temporal.

Todo o trabalho foi realizado com recurso ao programa estatístico *R* versão 3.0.1 (2013-05-16).

Palavras-Chave: Previsão, séries temporais, *Holt-Winters*, *ARIMA*, redes neurais e previsão da produção de eletricidade de base eólica.

Abstract

This research is concerned with using time series forecasting method to forecast the production of the electric power from wind energy in Cape Verde, particularly in Santiago Island. A special attention has been developed around this issue over the last few years. Forecasting the future, especially the temporal series behavior is something crucial in analysis and supporting decision-making. However, it continues to be a challenge for statistical and computation.

In this work it was used *Holt-Winters* model, *ARIMA* and *artificial neural nets*, RNAs radial basis functions (*RNAs-RBF*) and Perceptron of multiple layers (*RNAs-MLP*). The *Holt-Winters* model is an exponential forecasting model, known to deal with tendency elements and seasonality of a temporal series. The *ARIMA* model that contains only one variable, on the other hand, describes a variable behavior in terms of its past data. The neural nets have been serving as great tools in the application of temporal series forecasting. This research proposes a comparative analysis of these non-linear guide forecasting models in order to find out the model that best adapts to the temporal series.

All this work has been carried out resorting to statistical program R version 3.0.1 (2013-05-16).

Keywords: Forecasting, Temporal series, Holt-Winters, ARIMA, Neural Nets and Production electricity wind energy.

Índice

| | |
|--|-----|
| Nota Biográfica | ii |
| Agradecimentos | iii |
| Resumo | iv |
| Abstract..... | v |
| Índice de figuras | ix |
| Índice de tabelas | xii |
| 1. Introdução..... | 1 |
| 1.1. Objetivos da Dissertação | 2 |
| 1.2. Enquadramento do Tema..... | 2 |
| 2. Cabo Verde | 4 |
| 2.1. Potencial da Energia Elétrica em Cabo Verde | 5 |
| 2.2. Quadro Evolutivo da Produção da Energia Elétrica, 2007-2011 | 6 |
| 2.2.1. Percentagem de Produção por Ilha | 6 |
| 2.2.2. Potencial Eólico em Cabo Verde..... | 7 |
| 2.3. Santiago | 8 |
| 2.3.1. Potencial da Energia Elétrica na Ilha de Santiago | 9 |

| | |
|--|----|
| 2.3.2. Caraterização do Parque da Ilha de Santiago | 10 |
| 2.3.3. Potencial Eólico em Santiago | 11 |
| 3. Estado da Arte | 13 |
| 3.1. Importância da Previsão | 13 |
| 3.2. Horizontes de Previsão da Produção de Energia Eólica..... | 14 |
| 3.3. Modelos Físicos de Previsão da Energia Eólica..... | 15 |
| 3.4. Modelos Estatísticos e de Inteligência Computacional de Previsão da Produção da Eletricidade de Base Eólica | 19 |
| 3.5. Modelos Híbridos | 26 |
| 4. Previsão da Energia Eólica | 28 |
| 4.1. Relação entre a Velocidade do Vento e a Potência | 28 |
| 4.2. Previsão da Energia Eólica, Modelos Estatísticos..... | 30 |
| 4.2.1. Medidas Estatísticas Padrões (ME, MAE, MSE, RMSE e MAPE) | 31 |
| 4.2.2. Modelo de Suavização Exponencial de Holt-Winters (ES) | 33 |
| 4.2.3. Modelo ARIMA | 38 |
| 4.2.4. Redes Neurais Artificiais – RNAs..... | 42 |
| 4.2.4.1. Perceptron de Múltiplas Camadas (RNAs- MLP) | 42 |
| 4.2.4.2. Função de Base Radial (RNAs-RBF) | 55 |
| 4.2.5. Discussão dos Resultados Obtidos | 61 |

| | | |
|----|---------------------------------|----|
| 5. | Conclusões..... | 63 |
| 6. | Referências Bibliográficas..... | 66 |
| 7. | Anexo | 69 |

Índice de figuras

| | |
|--|----|
| Fig. 1: Mapa de Cabo Verde..... | 4 |
| Fig. 2: Matriz elétrica de Cabo Verde, ano 2011. Fonte: relatório e contas da ELECTRA, 2011. | 5 |
| Fig. 3: Percentagem de produção de energia elétrica em C. V., ano 2011, por ilha em kWh. | 6 |
| Fig. 4: Mapa da Ilha de Santiago. | 8 |
| Fig. 5: Produção da energia em Santiago ano 2008 a 2011 em kWh. | 10 |
| Fig. 6: Diagrama geral de modelos físicos [11]..... | 16 |
| Fig. 7: Exemplo de arquitetura da rede neural MPL..... | 22 |
| Fig. 8: Estrutura básica de uma rede neuronal RBF. Adaptado [26]. | 24 |
| Fig. 9: Combinação dos diferentes modelos de previsão da energia eólica [13]..... | 27 |
| Fig. 10: Curva de potência típica de um aerogerador [16]. | 28 |
| Fig. 11: Curva de potência que representa a produção em função da velocidade do vento do parque eólico da ilha de Santiago. | 29 |
| Fig. 12: Distribuição da frequência horária da potência do parque eólico da ilha de Santiago. | 35 |
| Fig. 13: Série da potência produzida de janeiro a dezembro de 2012. Os resultados são agregados à hora. Fonte: Elaborado pelo autor (2013) com base no banco de dados da empresa cabo-verdiana Cabeólica. | 36 |

| | |
|---|----|
| Fig. 14: Gráfico da potência observada e da potência prevista nas próximas 24 horas utilizando o modelo <i>Holt Winters</i> | 38 |
| Fig. 15: Gráficos da ACF e PACF..... | 40 |
| Fig. 16: Gráfico da potência observada e da potência prevista nas próximas 24 horas utilizando o modelo <i>ARIMA 4, 1, 1</i> | 40 |
| Fig. 17: Gráfico de teste de todas as redes MLP (outubro a dezembro de 2012). | 45 |
| Fig. 18: Gráfico do desempenho dos modelos de <i>RNs-MLP</i> (outubro a dezembro de 2012). | 45 |
| Fig. 19: Rede MLP (6,13,1) utilizada na modelação da série..... | 46 |
| Fig. 20: Gráfico de teste de todas as redes MLP (janeiro a março de 2012). | 47 |
| Fig. 21: Gráfico do desempenho dos modelos de <i>RNs-MLP</i> (janeiro a março de 2012). .. | 47 |
| Fig. 22: Rede MLP (6,13,1) utilizada na modelação da série..... | 48 |
| Fig. 23: Gráfico do teste de todas as <i>RNs-MLP</i> (junho a agosto de 2012). | 49 |
| Fig. 24: Gráfico do desempenho dos modelos de <i>RNs-MLP</i> , (junho a agosto). | 50 |
| Fig. 25: Rede MLP (3,7,1) utilizada na modelação da série..... | 51 |
| Fig. 26: Rede MLP (8,17,1) utilizada na modelação da série..... | 53 |
| Fig. 27: Gráfico dos erros obtidos na previsão da potência pelas 24 redes MLP. | 54 |
| Fig. 28: Gráfico dos erros obtidos na previsão da potência pelas 24 redes MLP | 54 |
| Fig. 29: Gráfico dos erros obtidos na previsão da potência pelas redes MLP..... | 55 |
| Fig. 30: Gráfico dos índices de desempenho da rede RBF. | 57 |

| | |
|---|----|
| Fig. 31: Gráfico dos erros obtidos na previsão da potência pelas redes RBF nas diferentes simulações..... | 57 |
| Fig. 32: Gráfico do treino da rede RBF. | 58 |
| Fig. 33: Gráfico do comportamento do modelo RBF, simulação 1. | 58 |
| Fig. 34: Gráfico do comportamento do modelo RBF, simulação 2. | 59 |
| Fig. 35: Gráfico do comportamento do modelo RBF, simulação 3. | 59 |
| Fig. 36: Gráfico do comportamento do modelo RBF, simulação 4. | 60 |
| Fig. 37: Gráfico do comportamento do modelo RBF, simulação 5. | 60 |
| Fig. 38: Gráfico do comportamento do modelo RBF, simulação 6. | 61 |

Índice de tabelas

| | |
|---|----|
| Tabela 1. Evolução da produção da energia elétrica, 2007-2011. | 6 |
| Tabela 2: Produção da energia elétrica (kWh), Santiago, ano 2011. | 9 |
| Tabela 3: Dados das turbinas do parque da ilha de Santiago. Fonte: projeto Cabeólica. | 11 |
| Tabela 4: Tabela da estatística descritiva da Potência. | 34 |
| Tabela 5: Teste de normalidade da variável potência. | 36 |
| Tabela 6: Valores dos parâmetros estimados utilizando o modelo <i>Holt Winters</i> , Alisamento Exponencial. | 37 |
| Tabela 7: Alguns modelos adequados utilizados na modelação das séries. | 39 |
| Tabela 8: Comparação dos erros entre Holt- Winters e ARIMA. | 41 |
| Tabela 9: Descrição das siglas. | 43 |
| Tabela 10: Modelação da série temporal utilizando Multi Layer Perceptrons. | 44 |
| Tabela 11: Modelação da série temporal utilizando Multi Layer Perceptrons. | 46 |
| Tabela 12: Modelação da série temporal utilizando Multi Layer Perceptrons. | 49 |
| Tabela 13: Erros obtidos na modelação das 24 redes MLP. | 52 |
| Tabela 14: índices de desempenho da rede RBF. | 56 |
| Tabela 15: Índice do desempenho de cada modelo utilizado na modelação das séries. | 61 |

1. Introdução

A energia elétrica em Cabo Verde é transformada a partir de diversas fontes de energia: sol (através de células fotoelétricas ou células fotovoltaicas), vento (através de aerogeradores) e combustíveis fósseis (geradores). Cabo Verde como é um país pobre e insular, depende muito da importação de combustível que para além de ser poluente é caro e sensível às flutuações dos preços de mercados. A inexistência de transporte de eletricidade entre as ilhas dificulta a gestão dos sistemas elétricos de energia como um único sistema, obrigando a que em cada ilha os sistemas elétricos de energia funcionem isolados. Tem-se assistido a problemas de instabilidade no fornecimento de eletricidade, sendo mesmo comum a existência de cortes de energia elétrica nas ilhas.

Sendo o vento um dos raros recursos naturais disponível em Cabo Verde, a produção da energia eólica é vista cada vez mais como uma opção viável de fonte de energia renovável, chegando mesmo a ser considerada uma das melhores. Para além de ser um recurso energético natural, este deve ser explorado e utilizado no máximo, para benefício do país e do futuro dos cabo-verdianos.

Este trabalho é realizado no âmbito do Mestrado em Análise de Dados e Sistemas de Apoio à Decisão, na Faculdade de Economia da Universidade do Porto - FEP e pretende utilizar modelos de previsão para a produção do total de energia eólica em Cabo Verde, particularmente na ilha de Santiago. Para a sua realização serão utilizadas as seguintes metodologias:

- Recolha e tratamento de dados;
- Escolha de modelos;
- Definição do tipo de previsão a ser feito (período de previsão): Curto Prazo – Horizonte temporal até vinte e quatro horas;
- Utilização dos dados de teste para avaliar o desempenho dos diferentes modelos de previsão;
- Análise dos resultados/conclusões.

Os dados de treino e de teste foram relativos aos meses de janeiro a dezembro de 2012. Serão utilizados Modelos Estatísticos e de Inteligência Computacional, que permitem relacionar os valores históricos, velocidade do vento, velocidade do vento normalizada, direção do vento, temperatura, pressão e massa do ar com as previsões das potências no horizonte temporal a curto prazo.

1.1. Objetivos da Dissertação

São apresentados os objetivos que norteiam este trabalho:

- Utilizar modelos de previsão da potência da produção da energia eólica no parque da ilha de Santiago a curto prazo;
- Utilizar vários previsores e estimar os seus desempenhos;
- Analisar os resultados das previsões obtidas;
- Comparar os resultados obtidos entre os diferentes modelos utilizados nos estudos das séries temporais.

1.2. Enquadramento do Tema

Denomina-se energia eólica a energia cinética contida nas massas de ar em movimento (vento). Seu aproveitamento ocorre por meio da conversão da energia cinética de translação em energia cinética de rotação, com o emprego de turbinas eólicas, também denominadas aerogeradores, para a geração de eletricidade.

As primeiras tentativas para a geração de eletricidade a partir da eólica surgiram no final do século XIX, mas somente um século depois, com a crise internacional do petróleo (década

de 1970), é que houve interesse e investimentos suficientes para viabilizar o desenvolvimento e aplicação de equipamentos em escala comercial.

A primeira turbina eólica comercial ligada à rede elétrica pública foi instalada em 1976, na Dinamarca. Atualmente existem mais de 30 mil turbinas eólicas em operação no mundo. Em 1991, a Associação Europeia de Energia Eólica estabeleceu como metas a instalação de 4.000 MW de energia eólica na Europa até o ano 2000 e 11.500 MW até o ano 2005. Essas e outras metas foram cumpridas muito antes do esperado (4.000 MW em 1996, 11.500 MW em 2001) [1].

Estima-se que em 2020 o mundo terá 12% da energia gerada pelo vento, com uma capacidade instalada demais de 1.200 GW [WINDPOWER; EWEA; GREENPEACE, 2003; WIND FORCE, 2003].

A energia eólica em Cabo Verde está enquadrada no desenvolvimento estratégico do Governo em aumentar a participação das energias renováveis no arquipélago para 25% em 2011 e para 50% até 2020. Neste sentido foram instalados quatro parques eólicos nas ilhas de São Vicente, Boa Vista, Sal e Santiago, tendo capacidades para produzir até 25,5 MW de energia renovável e beneficiar cerca de 85% dos 491.575 habitantes [2].

As ferramentas de previsão da produção são importantes no apoio ao planeamento e à gestão, de modo a permitir por um lado a otimização da operação do sistema elétrico e por outro garantindo a segurança no abastecimento, dado que o recurso vento apresenta uma característica variável.

2. Cabo Verde

O arquipélago de Cabo Verde, com capital na Cidade da Praia (ilha de Santiago) tem uma superfície de 4.033 km² e uma população residente de 491.575 indivíduos [2]. Situa-se a oeste do Continente Africano, entre os paralelos 15 e 17 graus de latitude norte e as longitudes de 22° 41' e 25° 22' Oeste de Greenwich. Localizadas em pleno Oceano Atlântico, ao largo do Senegal e da Mauritânia, está a uma distância de cerca de 455 km da Costa de África (Fig. 1). Pelo facto de ficar muito próximo do Cabo Verde recebeu o nome de Arquipélago de Cabo Verde. É constituído por 10 ilhas (Santo Antão, São Vicente, Santa Luzia, São Nicolau, Sal, Boavista, Maio, Santiago, Fogo e Brava), sendo uma desabitada, a de Santa Luzia e vários ilhéus (Branco, Raso, Grande, Luís Carneiro, Cima os ilhéus do Rombo ou Secos, ilhéu dos Pássaros e ilhéu de Santa Maria), dispondo de um espaço marítimo exclusivo de mais de 600.000 km².

As ilhas dividem-se em dois grupos, consoante a sua posição geográfica relativamente aos ventos dominantes: Barlavento, do lado que sopra o vento, (Santo Antão, São Vicente, Santa Luzia, São Nicolau e Sal, Boavista e os ilhéus os ilhéus Branco e Raso) e Sotavento, do lado oposto ao vento, (Brava, Fogo, Santiago, Maio e os ilhéus Grande, Luís Carneiro, Cima e os ilhéus do Rombo ou Secos e de Santa Maria) [3].

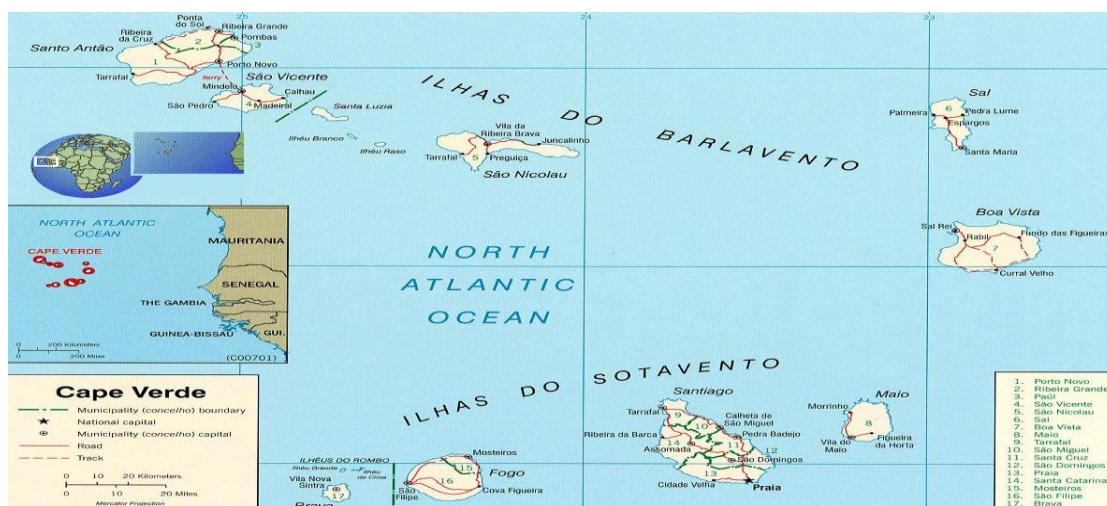


Fig. 1: Mapa de Cabo Verde.

2.1. Potencial da Energia Elétrica em Cabo Verde

Cabo verde é um dos países que não dispõe de explorações de combustíveis fósseis, por inexistência desse recurso ou por ser inviável a sua exploração. Sendo assim, a dependência externa deste recurso não renovável é uma das principais dificuldades do setor elétrico. A potência instalada no parque produtor da ELECTRA totalizava, no final do ano de 2011, 94.331 kW repartidos pelas centrais a diesel 84.731 kW (90 %), centrais eólicas 2.100 kW (2%) e centrais solares fotovoltaicas 7.500 kW (8%).

Segundo o relatório e contas de 2011 da empresa de abastecimento de água e energia, ELECTRA, SARL, da eletricidade produzida durante 2011, mais de 90% foi a partir de combustíveis fósseis (fig. 2).

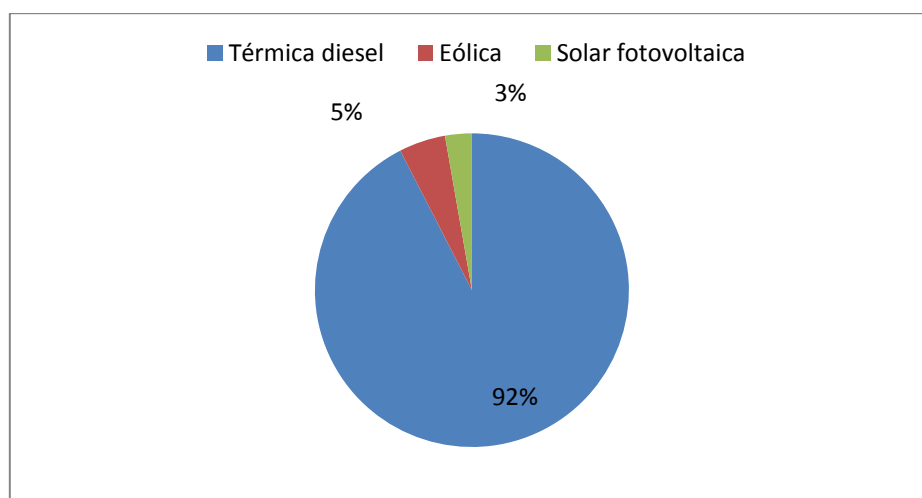


Fig. 2: Matriz elétrica de Cabo Verde, ano 2011. Fonte: relatório e contas da ELECTRA, 2011.

2.2. Quadro Evolutivo da Produção da Energia Elétrica, 2007-2011

Segundo o quadro da evolução da produção da energia elétrica, (relatório da ELECTRA 2011, tabela 1), no ano de 2011 foram produzidas 325.421 MWh de energia elétrica, sendo 92% de origem térmica, 5% de origem eólica e 3% solar.

Tabela 1. Evolução da produção da energia elétrica, 2007-2011.

| Grupos de Tecnologias | 2007 | 2008 | 2009 | 2010 | 2011 |
|-----------------------|--------|--------|--------|--------|--------|
| Térmica diesel | 260465 | 279645 | 290273 | 314315 | 300860 |
| Eólica | 6869 | 5510 | 4661 | 1992 | 15605 |
| Vapor | 1185 | 640 | 0 | 0 | 0 |
| Solar Fotovoltaica | 0 | 0 | 0 | 2105 | 8956 |
| TOTAL | 268518 | 285795 | 294934 | 318413 | 325421 |

A produção de energia térmica diesel entre 2010 e 2011 teve uma diminuição de 4,3%, enquanto a eólica mostrou aumentos significativos, juntamente com a geração solar cujas percentagens foram de 683% e 325% respetivamente, conforme tabela 1 [4].

2.2.1. Percentagem de Produção por Ilha

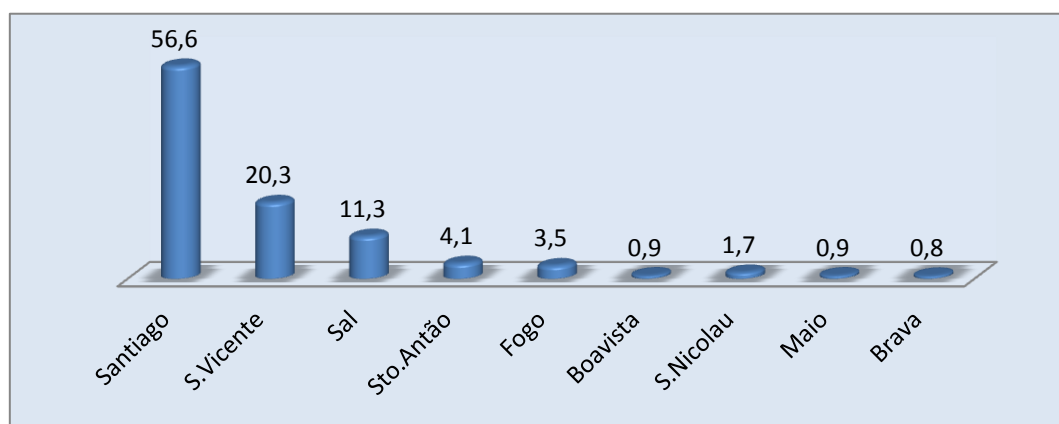


Fig. 3: Percentagem de produção de energia elétrica em C. V., ano 2011, por ilha em kWh.

Entre os diferentes centros eletroprodutores podemos constatar segundo o relatório, que as ilhas de Santiago, São Vicente e Sal totalizam um valor de produção de aproximadamente 88% do total do país. A fig. 3 mostra esta distribuição, destacando-se a ilha de Santiago com mais de 56% de produção no final de 2011.

2.2.2. Potencial Eólico em Cabo Verde

Devido à sua localização geográfica, a meio do Oceano Atlântico, Cabo Verde dispõe de recursos eólicos extremamente favoráveis. Sofre influência de ventos alísios, podendo estes ser utilizados para a produção de energia elétrica. Foi neste sentido que, em 1994, a ELECTRA procedeu à instalação de três parques eólicos com capacidade de 2x300 kW na ilha do Sal e 3x300 kW nas ilhas de S. Vicente e Santiago. Estes parques são compostos por aerogeradores da Nordtank de 300 kW, com regulação “Stall” [5].

Nos finais de 2011, a Vestas (multinacional ligada à energia eólica) assinou um contrato para o fornecimento de geradores eólicos que totalizassem a capacidade de 25,5 MW para Cabo Verde, especificamente para o projeto Cabeólico, que incluiu a construção de 4 parques nas ilhas de Santiago, Sal, São Vicente e Boavista. A Vestas instalou um total de trinta máquinas de 850 kW de potência uma parceria Público / Privada entre o Governo de Cabo Verde, ELECTRA & InfraCo.

Em 2012, Cabo Verde, aparece destacado como país com grande aproveitamento do vento para a eletricidade, com uma produção de 8 MW de potência instalada. Os outros países apostam em turbinas eólicas também para outros fins. "Apesar do grande avanço da China, Cabo Verde é o país que mais cresceu, em média, na utilização de energia eólica", realça o Conselho Mundial de Energia Eólica, facto considerado pelo site especializado em fontes alternativas, *CleanTechnical*, como "fantástico para um país com um PIB tão baixo".

Ainda segundo o mesmo relatório, o arquipélago também lidera o ranking de novas unidades instaladas em 2011 (10 MW, a partir das 11 turbinas do Monte São Filipe, na

Cidade da Praia, ilha de Santiago), seguido das Honduras, que instalou 6 MW, e da China que investiu 2 MW no ano de 2010). [relatório do Conselho Mundial de Energia Eólica (GWEC, sigla em inglês), referente a 2011].

2.3. Santiago

Santiago é a maior ilha de Cabo Verde, com cerca de 273.919 habitantes, cerca de 56% da população de Cabo Verde (censo de 2010), pertence ao grupo de Sotavento. Estende-se por cerca de 75 km de comprimento, no sentido Norte-Sul e cerca de 35 km de largura, no sentido Leste-Oeste e com uma área de 991m². Dista cerca de 50 km em linha reta da ilha do Fogo, a oeste, e 25 km da ilha do Maio, a leste. Administrativamente, está dividida em 9 concelhos. A cidade da Praia é ao mesmo tempo a capital do país e a cidade mais populosa com 131.602 habitantes (Extraído do original: http://pt.wikipedia.org/wiki/Cabo_Verde).



Fig. 4: Mapa da Ilha de Santiago.

2.3.1. Potencial da Energia Elétrica na Ilha de Santiago

A ilha de Santiago registou um crescimento maior que as demais ilhas com produções bem expressivas (Fig. 3), acima dos 56%. Na cidade da Praia, ilha de Santiago, que é contemplada com as centrais da Praia e do Palmarejo a produção é acima dos 48% e no interior estão instaladas as centrais de Assomada - Santa Catarina, Ribeira da Barca, Tarrafal e Santa Cruz totalizam uma produção de cerca de 8,4% (relatório da Eletra de 2011, tabela 2).

Tabela 2: Produção da energia elétrica (kWh), Santiago, ano 2011.

| Ilha | Central | Diesel | Eólica | Solar | Total | Repartição percentual |
|--------------------------------|---------|--------------------|------------------|------------------|--------------------|-----------------------|
| Cidade da Praia | | 6.268.750 | 7.727.170 | | 13.995.920 | |
| Palmarejo | | 136.014.115 | | 6.893.418 | 142.907.532 | |
| Total Praia | | 142.282.865 | 7.727.170 | 6.893.418 | 156.903.452 | 48,2% |
| Assomada (Sta Catarina) | | 13.646.807 | | | 13.646.807 | |
| Rib ^a da Barca | | 0 | | | 0 | |
| Tarrafal Santiago | | 5.421.132 | | | 5.421.132 | |
| S.Cruz | | 8.285.633 | | | 8.285.633 | |
| Total Interior Santiago | | 27.353.572 | | | 27.353.572 | 8,4% |
| Santiago | | 169.636.437 | 7.727.170 | | 184.257.024 | 56,6% |

A produção da energia elétrica na ilha provém de três fontes: geração a diesel, a eólica e a solar fotovoltaica, com maior produção proveniente da térmica diesel.

No ano 2011 a ponta máxima produzida foi de 36093 kWh de energia elétrica, sendo 92,1% de origem térmica, 4,2% eólica e 3,7% solar.

O gráfico que se segue mostra a evolução da produção da ponta máxima de 2008 a 2011:

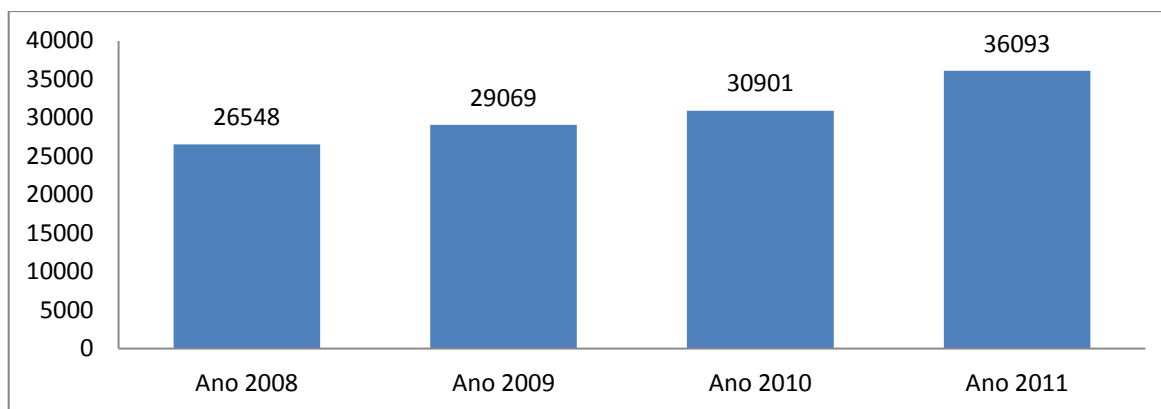


Fig. 5: Produção da energia em Santiago ano 2008 a 2011 em kWh.

Ao longo dos anos verificou-se um crescimento razoável, passando de 26548 kWh em 2008 para 36093 kWh em 2011. Um aumento de aproximadamente 36% (Fig. 5).

2.3.2. Caraterização do Parque da Ilha de Santiago

O parque em estudo fica situado na zona de Monte São Filipe ou ilhéu de São Filipe, cidade da Praia, ilha de Santiago. Trata-se de uma achada de altitude média que varia entre 170 e 300 metros. As superfícies deste planalto apresentam uma ligeira inclinação para o litoral Este e a área tem uma média anual de velocidade do vento de 8,4 m/s e a capacidade instalada é de 10 MW com 11 turbinas, de altura de 55 metros de altura, com potências de 850 kW e com a produção anual de 30,1 a 40,3 GWh/ano. A taxa de penetração é de 17% (tabela de produção anual de energia do projeto Cabeólica/Ministério do Ambiente Desenvolvimento Rural e Recursos Marinhos, 2009).

Cada aerogerador é constituído por uma torre de aço, de forma cónica tubular, com diâmetro de rotor de 26 m e altura de eixo de rotação de 44 m. No topo encontra-se instalada a cabina que aloja o sistema de transmissão, o gerador e a quase totalidade dos sistemas auxiliares e de segurança. A entrada do sistema de transmissão é feita através do veio principal rigidamente ligado ao cubo do rotor, constituído por três pás – em fibra de vidro e de poliéster – rodando, em condições de operação normais. A cabina é orientável,

rodando em torno de um eixo vertical, de forma a posicionar-se no azimuth do vento dominante [6]).

Estes aerogeradores eólicos funcionam com um nível de tensão de cerca de 690 V e cada um tem um transformador localizado ao lado da sua base da turbina para subir a tensão da turbina eólica para 20 kV [6].

O quadro que se segue caracteriza os dados das turbinas instaladas no parque eólico da ilha de Santiago:

Tabela 3: Dados das turbinas do parque da ilha de Santiago. Fonte: projeto Cabeólica.

| Turbinas | | | | | | |
|----------|------------|--------|----------|---------------|-------------------------|--------------|
| Nº | Fabricante | Modelo | Potência | Potênc. total | Potênc. anual (GWh/ano) | % Penetração |
| 11 | Vestas | V 52 | 850 kW | 10 MW | 30,1 – 40,3 | 17 |

2.3.3. Potencial Eólico em Santiago

A ilha de Santiago é uma das menos áridas do Arquipélago. Quase todo ano está sob a influência dos ventos alísios do hemisfério norte (Relatório Energias Renováveis, 2011). Sendo assim as possibilidades de exploração de energia eólica são excelentes.

A análise do regime de ventos reporta-se o período de 1988 a 2010 é a seguinte: os ventos dominantes da ilha de Santiago, como no restante do arquipélago são do noroeste, atingindo velocidades médias na ordem dos 5,3 m/s. O regime sazonal de ventos é dominado pela presença dos ventos alísios (ventos do quadrante noroeste), que sopra predominantemente entre novembro a maio em toda a faixa litoral oriental (Relatório Energias Renováveis, 2011).

A Divisão de Energias Renováveis (DER) do Ministério do Desenvolvimento Rural de Cabo Verde (MDR) realizou, em 1977, medições de vento em 25 localidades para analisar

o regime de vento. Os locais caracterizados como sendo viáveis para a instalação dos equipamentos com uma potencial eólica satisfatória na ilha de Santiago foram: Achada São Filipe, Pensamento, Lem Duque, Grancha S. Filipe, Trindade, Santa Cruz, João Varela São Pedro, Salineiro, Ribeira Flamengos, Ribeirão Manel, Ponta Furna, Portete e Pensamento (Relatório Energias Renováveis, 1984).

De acordo com esse relatório, a conclusão que se chegou é de que a ilha é “rica” em potencial eólico e que seu aproveitamento deve ser melhor estudado com novas medições do vento. Conclusão partilhada por outros autores, como Fonseca, que estima um grande potencial eólico inexplorado até então na ilha. A falta de estudos específicos, impossibilita afirmar categoricamente qual o potencial dessa ilha.

Na ilha de Santiago, no Monte de São Filipe, foi instalado um parque eólico de 10 MW de potência, conforme a figura em anexo. (Cabeólica Wind Farm Project, 2010).

O parque eólico está conectado à rede elétrica existente na ilha e os cabos de transmissão são totalmente subterrâneos. As turbinas operam a velocidades do vento entre 4m/s e 25 m/s (aproximadamente). Se forem expostos a velocidades do vento superior a 28 m/s ao nível da altura das torres, as turbinas desligam para autoproteção. Estas condições do vento são consideradas raras na zona da instalação dos aerogeradores.

3. Estado da Arte

3.1. Importância da Previsão

A previsão é um desperdício se não for integrada de modo a que a informação seja útil na operação do sistema e nos mercados elétricos. Se os operadores do sistema não têm confiança nas previsões, ou não é disponibilizada de uma forma que lhe seja útil utilizar, são tentados a operar o sistema num modo mais conservador aumentando assim os custos da integração da energia eólica [7].

Mas também a sua importância deve-se aos seguintes fatos:

Ajuda na gestão do congestionamento – Identificação de possíveis instabilidades na rede como desconexões de equipamentos de proteção devido a ecos de tensão ou perturbações na frequência [8];

Estimação da reserva - Devido à natureza volátil do vento, a produção eólica não oferece garantia de potência, o que obriga dispor de reserva adicional de outras fontes de produção para fazer frente a uma possível ausência da produção eólica [8];

Estimação da produção horária - Os sistemas de previsão da energia eólica permitem estimar a produção horária de um parque ou conjunto de parques eólicos para um horizonte de vários dias. Esta previsão faz diminuir os desvios nos programas de produção e minimiza o impacto na operação do sistema, fazendo com que esta seja mais fiável e segura.

Dispor de bons sistemas de previsão permite aumentar o nível de penetração da energia eólica no sistema elétrico, sem por em causa a segurança e fiabilidade na operação do sistema.

3.2. Horizontes de Previsão da Produção de Energia Eólica

Para avaliar o potencial energético de um determinado lugar a fim de determinar as condições para a melhor produção da energia elétrica dos aerogeradores, é necessário efetuar a medição das condições meteorológicas: velocidade e direção dos ventos, pressão atmosférica, temperatura, massa volúmica do ar e densidade do ar. Isso permitirá saber com o máximo de precisão possível, qual o regime local do vento e consequentemente o recurso eólico disponível. Os dados obtidos poderão ser utilizados pelos modelos de previsão da produção de energia eólica durante a fase de modelação dos mesmos.

Os modelos de previsão são caracterizados pelos seus horizontes temporais, como sendo: de muito curto prazo, curto prazo, médio prazo e de longo prazo e pela agregação/categorias em, modelos Físicos, modelos Estatísticos e modelos Híbridos.

Previsões de Muito Curto Prazo – Os sistemas operam para horizontes temporais de segundos até algumas horas. Estão associados a situações de controlo dos aerogeradores, de definição da garantia de segurança de abastecimento por parte do operador da rede elétrica e de gestão de unidades convencionais rápidas que participam nos serviços de sistema (muito usual em sistemas isolados, como por exemplo ilhas) [9];

Previsões de Curto Prazo – Este tipo de previsões estão normalmente definidas através do intervalo de previsão nas próximas horas, entre os 30 minutos às 72 horas. As previsões de curto prazo estão normalmente associadas ao apoio na gestão das centrais elétricas convencionais e para programar os planos de manutenção dos aerogeradores para as horas em que a previsão da produção eólica é baixa ou nula;

Previsões de Médio Prazo – Este tipo de previsões estão normalmente definidas através do intervalo de previsão nos próximos dias, entre os 7 ou 10 dias. As previsões de médio prazo estão normalmente associadas ao apoio na gestão das centrais elétricas convencionais e

para programar os planos de manutenção dos aerogeradores para os dias em que a previsão da produção eólica é baixa ou nula;

Previsões de Longo Prazo – Os intervalos de previsões podem variar entre vários dias ou até podem chegar a meses ou anos. Previsões de longo prazo são úteis na seleção de locais para a localização de parques eólicos.

No trabalho irão ser abordadas as previsões de curto prazo em que os horizontes temporais podem variar de 60 minutos às 24 horas.

Os modelos de previsão da energia eólica a curto prazo utilizam diferentes informações de entrada, tais como: informações da previsão do tempo, estado do tempo no local dos parques, informações do estado de funcionamento em tempo real dos parques eólicos fornecidos pelo sistema SCADA (*Supervisory Control and Data Acquisition*), que permite o controlo remoto e monitorização individual de cada aerogerador e do parque eólico, quer a partir do computador central ou de um computador pessoal remoto) e informações adicionais sobre características dos parques eólicos e a orografia¹ dos terrenos perto dos parques. Estes modelos de previsão da energia eólica podem ser classificados em modelos físicos, estatísticos e ou de inteligência computacional.

3.3. Modelos Físicos de Previsão da Energia Eólica

Modelos que fornecem previsões de velocidade e direção do vento com base em informações meteorológicas (pressão, temperatura, humidade e entre outras) e físicas (características do terreno, altitude, rugosidade, obstáculos e entre outras).

¹Descrição do relevo de uma região

- **O modelo NWP**

NWP (*Numerical Weather Prediction*) é um modelo de previsão numérica que se baseia num conjunto de equações que têm forma de equações diferenciais parciais que traduzem as leis físicas que descrevem o comportamento hidrodinâmico da atmosfera. Utiliza modelos matemáticos para prever o tempo com base nas condições climáticas atuais. Podem ser utilizados para gerar previsões de curto prazo de tempo ou previsões de longo prazo, sendo este último amplamente aplicado para entender e projetar mudanças climáticas.

É de salientar que os modelos de previsão numérica são capazes de fornecer valores de previsão à escala continental, para um horizonte que varia de 3 a 72 horas [7]. Eles têm em consideração o fluxo de ventos dentro e na periferia do parque eólico.

O esquema que se segue representa o modo de funcionamento deste tipo de modelo:

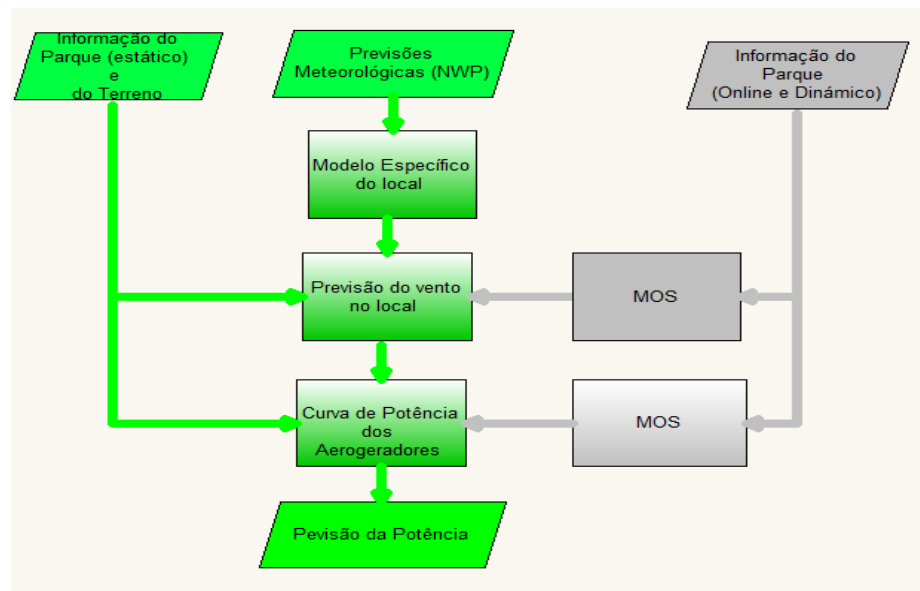


Fig. 6: Diagrama geral de modelos físicos [11].

Para melhorar os seus desempenhos, por vezes, os modelos físicos recorrem à agregação de modelos estatísticos, que operam com base na informação adquirida no parque em tempo

real, como demonstra a Fig. 6. Pois, os modelos apresentam erros sistemáticos que podem ser corrigidos com a aplicação de modelos estatísticos lineares ou não, a partir das estatísticas acumuladas de um certo modelo de previsão. Esses modelos corretivos à posteriori têm o nome genérico de *Model Output Statistics* (MOS). Os modelos NWP incorporam internamente um modelo MOS, que adaptam as saídas dos modelos físicos com as medições efetuadas em tempo real.

Os modelos físicos recorrem às condições físicas atmosféricas para conseguirem melhores previsões possíveis da velocidade do vento no local, e o modelo MOS é introduzido por forma a reduzir o erro. Para converter as previsões da velocidade em potência são aplicados modelos de curva de potência que têm como princípio de funcionamento a transposição de valores de velocidade de vento em valores de potência elétrica, recorrendo à curva de potência característica de cada aerogerador. A curva de potência pode ser fornecida pelo fabricante ou pode ser identificado por processamentos históricos de séries temporais com potência produzida pela máquina em função das características do vento (velocidade, direção, e entre outras). Quando a curva de potência é disponibilizada pelo fabricante, normalmente surge um erro na previsão da potência do parque. Este erro surge uma vez que a curva de potência não é dependente apenas das características da máquina, mas também da orografia do terreno e do próprio desempenho da máquina, consequência da parametrização dos coeficientes de controlo de cada turbina [11].

Requerem muitos dados e ainda demandam grande esforço computacional e por este motivo a sua utilização na previsão de curto prazo é bastante limitado. Segundo [12] é adequado para horizonte de previsões superiores a 6 horas.

- **Metodologia Wasp**

WAsP (*Wind Atlas Analysis and Application Program*) é um programa de modelação simplificada dos escoamentos atmosféricos desenvolvido pela divisão de Energia Eólica da Universidade Técnica RISø, na Dinamarca. É um modelo físico que calcula estimativas de

produção de energia elétrica em parques eólicos, estimativas de regimes locais de ventos e estimativas de produção de aerogeradores. Estas previsões são baseadas em medições do vento em estações meteorológicas que estejam na proximidade ou na futura área de implementação de um parque eólico.

Calcula a potência prevista usando a curva de potência [10].

Como qualquer outro modelo o WAsP necessita de informações sobre condições meteorológicas iniciais que são fornecidas por um modelo de escala superior (ou por observações), a fim de poder obter simulações saídas do WRF (um modelo de mesoscala) que serão usadas como dados de entrada no modelo de microscala. Este modelo irá considerar as simulações (ou observações) num dado ponto e construir um regime regional de ventos (ou atlas de vento), que consiste nas observações/simulações (dados de entrada) limpos dos efeitos locais do terreno no ponto [12].

- **Modelo MM5**

MM5 (*The Fifth-Generation NCAR / Penn State Mesoscale Model*) é um modelo numérico de mesoscala utilizado para simular e prever o comportamento dos padrões atmosféricos à escala regional. O código do modelo é de acesso livre e continua sendo aperfeiçoado por diversos utilizadores em universidades e institutos de investigação. Utiliza coordenadas sigma, que acompanham o terreno, o que o torna capaz de ser utilizado operacionalmente para fins de previsão atmosférica.

Este modelo faz em geral recurso a meios informáticos poderosos em termos de memória e velocidade de processador. Este modelo resolve um sistema complexo de equações matemáticas baseadas em leis físicas de modo a prever o estado futuro da atmosfera partindo de condições iniciais específicas.

Os modelos de previsão numérica do tempo podem ser classificados segundo a sua escala espacial dentro da qual a previsão é feita:

Global, com resolução da ordem de 200 km, visa identificar o comportamento geral da atmosfera sobre uma área extensa, estes modelos são apenas capazes de identificar fenómenos meteorológicos de larga escala (ditos sinóticos);

Regional ou local, com resolução espacial de alguns metros a cerca de 50 km, visa identificar com maior detalhe o comportamento da atmosfera sobre uma região específica, sendo capazes de identificar fenómenos meteorológicos de pequena escala (ditos de mesoscala) [12].

3.4. Modelos Estatísticos e de Inteligência Computacional de Previsão da Produção da Eletricidade de Base Eólica

Estes modelos não consideram qualquer efeito físico, ou seja não usam tanta informação específica dos locais dos parques eólicos (orografia do terreno, características dos aerogeradores) como os modelos físicos. Simplesmente relacionam os valores históricos de potência registada com variáveis como, a velocidade e direção do vento, evitando desta forma elevados tempos de processamento em comparação com os modelos físicos.

- **Modelo Estatístico -MOS**

MOS (*Model output statistics*) é um modelo estatístico, considerado importante na previsão do tempo [13].

De um modo geral, os resultados dos modelos numéricos representam melhor os padrões da atmosfera em altitude sendo mais débeis quando tentam prever as variações locais à superfície. Por outro lado, os modelos estatísticos puros conseguem representar a previsão do tempo local, mas perdem qualidade para além de seis horas de previsão. O MOS é uma

técnica que combina estes dois métodos, usando as previsões numéricas com base na física da atmosfera para prever os padrões climáticos de grande escala e, as equações de regressão estatística resultantes do pós-processamento para esclarecer detalhes do tempo à superfície. A precisão dos resultados é geralmente muito melhor do que utilizar apenas um modelo estatístico ou numérico puro [13].

A deficiência principal deste método estatístico relaciona-se com a existência de um elevado número de dados disponíveis das saídas de uma mesma versão de modelo numérico para permitir um melhor desempenho das equações de previsão.

Os modelos de previsão estatísticos têm por base equações de regressões lineares múltiplas, onde a variável dependente é expressa em função das quantidades que contribuem para a sua variação.

- **Modelo estatístico de séries temporais**

É um modelo de previsão com base em séries temporais que gera valores futuros baseado exclusivamente em dados históricos. Esta é a abordagem proposta em modelos incluindo *ARMA*, *ARIMA* ou de *Box-Jenkins* [14], que é útil para a previsão de certos processos industriais, e no contexto da previsão da produção de eletricidade de base eólica. Fornece resultados razoáveis em horizontes temporais de até 6 horas [15].

No entanto, a utilização de modelos de previsões baseados em séries temporais clássicas não é muito propício para previsão de horizontes superiores a 10 horas.

Também é utilizada a função *auto-arima* que usa o algoritmo de “*Hyndman Khandakar algorithm*” para modelação de *ARIMA(p,d,q)* automático, de autocorrelação-FAC e da função de autocorrelação parcial-FACP estimadas para a série estacionarizada que segue o seguinte algoritmo:

1. O número de diferenças d é determinado usando testes $KPSS^2$ repetidas;

² **Kwiatkowski-Phillips-Schmidt-Shin** (KPSS) testes usados para testar uma hipótese nula de que uma dada série é estacionária em torno de uma tendência determinista.

2. Os valores de p e q são então escolhidos minimizando o $AICc$ após serem feitas diferenças dos dados d -vezes. Em vez de considerar cada combinação possível de p e q , o algoritmo usa uma pesquisa “*stepwise*” para todo o espaço do modelo.

a) O melhor modelo (com $AICc$ mais baixo) é selecionado da lista a seguir:

$ARIMA(1, d, 1)$;

$ARIMA(0, d, 1)$;

$ARIMA(1, d, 5)$.

Se $d=0$, em seguida, a constante c está incluída; Se $d \geq 1$, em seguida, a constante c é definida como zero. Isso é chamado de “modelo atual”.

b) São consideradas as seguintes variações no “modelo atual”:

- Variar p e/ou q do modelo atual por ± 1 ;
- Incluir/excluir c do modelo atual.

c) Repete o passo 2 até que o $AICc$ inferior seja encontrado.

Os argumentos de *auto.arima* () fornecem muitas variações sobre o algoritmo. O que está descrito acima é o comportamento padrão [17].

Este modelo inclui três componentes: autorregressivo (AR), integrador (I) e a média móvel (MA). No modelo autorregressivo (AR) o valor presente da série temporal X_t é expresso linearmente em termos dos valores anteriores da série e da perturbação aleatória a_t relativa ao instante t . A ordem deste processo depende do valor mais antigo sobre o qual a regressão se baseia [19].

Há outras abordagens alternativas para o modelo clássico de séries temporais, que prevê a potência de energia elétrica através de parques eólicos, o uso de modelos baseados em redes neurais:

- **Modelos baseados em redes neurais**

As redes neurais são soluções interessantes devido à sua flexibilidade de utilização numa multiplicidade de problemas. Como um exemplo, em [9] são utilizadas redes neurais recorrentes com múltiplas camadas dedicadas a horizontes de previsão diferentes de até 24

horas. Também foram utilizadas redes neuronais multicamadas para a predição da potência até 1 hora para o uso no controlo de um parque eólico, utilizando como entrada valores recentes da série temporal do vento [10]. As redes neuronais geralmente proporcionam bons resultados a curto prazo, tornando-se necessário ter outras fontes de informação, tais como a previsão meteorológica obtida com modelos de previsão atmosférico, para calcular com a precisão suficiente para potências em horizontes de predição de 4-6 horas.

As redes neuronais artificiais são sistemas paralelos distribuídos, compostos por unidades de processamento simples denominados neurónios e dispostas em uma ou mais camadas interligadas por um grande número de conexões (*sinapses*) geralmente unidireccionais e com pesos para ponderar as entradas recebidas por cada neurónio [16].

A arquitetura mais usual é a rede neural *Multi Layer Perceptron (MLP)* com três camadas, conforme ilustrado na Fig. 7. De forma análoga ao que acontece no cérebro humano, onde as *sinapses* são reforçadas ou enfraquecidas, os pesos nas conexões entre as camadas são ajustados durante o processo de aprendizagem da rede neural [16].

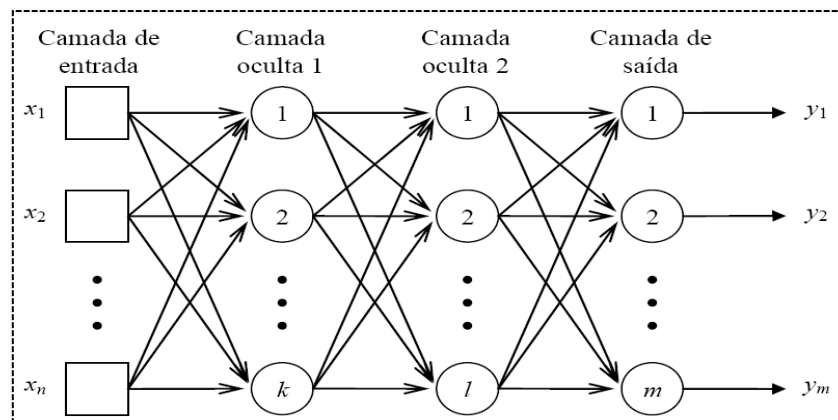


Fig. 7: Exemplo de arquitetura da rede neural MPL.

Uma RNA é composta por um certo número de neurónios artificiais organizados em diferentes camadas, (conferir Fig. 7). A primeira camada, chamada usualmente de camada de entrada, serve apenas para propiciar a entrada dos dados, não executando qualquer computação. É nas camadas intermediárias ou ocultas e na camada de saída, que ocorre o

efetivo processamento dos dados. A definição do número de camadas escondidas/ocultas e do número de neurónios em cada camada é feita de forma empírica e depende da distribuição dos padrões de entrada, da quantidade de ruído presente nos exemplos de treino e da complexidade da função a ser aprendida.

Segundo [22] a questão principal para operacionalizar uma *RNA* é como obter os vetores de pesos (w), que correspondem às sinapses que se estabelecem entre os neurónios biológicos. No cérebro humano as sinapses são estabelecidas, e fortalecidas ou não, através de um processo de aprendizado. Para uma *RNA* precisa-se também constituir um processo de aprendizagem para a determinação dos pesos que ligam os neurónios artificiais.

Basicamente existem dois modelos de aprendizagem [25]:

i) Aprendizagem supervisionada: necessita dos valores de entrada e suas respectivas saídas esperadas. A cada iteração desse conjunto de entradas e saídas é encontrado um valor de erro, utilizado para corrigir os pesos sinápticos e níveis de *bias*;

ii) Aprendizagem não-supervisionada: somente os valores de entrada são esperados. Para este caso a rede tenta descobrir padrões nos valores de entrada através de identificação de regularidades estatísticas. Dessa forma a rede neural automaticamente adapta os pesos sinápticos para representar e agrupar os padrões de entrada com características semelhantes. Também tem-se as *RNAs-RBF*, como podemos constatar na Fig. 8, que são redes neuronais que têm uma camada de entrada, uma camada oculta e uma camada de saída. Os neurónios na camada escondida contêm as funções de transferência de Gauss, cujas saídas são inversamente proporcionais à distância entre o centro do neurónio.

Esse tipo de rede foi originalmente desenvolvido para interpolação de dados em espaços multidimensionais [26]. Esse problema pode ser formulado pela seguinte equação:

$$y_i = F(U_i), \quad (3.1),$$

onde y_i é um conjunto de escalares e u_i é um conjunto de vetores. Desta forma, se a função $F(.)$ for definida analiticamente ela pode mapear valores do vetor u_i que não pertenciam ao conjunto original de treino, no conjunto de escalares y_i . Para que a função $F(u_i)$ seja uma solução analítica ela pode ser definida da seguinte forma:

$$F(U) = f(x) = \sum_{i=1}^n w_i \varphi(||u - u_i||^2) \quad (3.2),$$

onde w_i é um coeficiente (o peso da conexão neste caso) e $\varphi(||u - u_i||^2)$ representa uma função radial. Existindo n funções radiais, o centro de cada função é dado por u_i . Desta forma, $||u - u_i||$ representa a distância Euclidiana entre o centro de cada função radial (u_i) e do vetor de entrada (u).

A estrutura básica de uma rede neuronal RBF é ilustrada na figura que se segue:

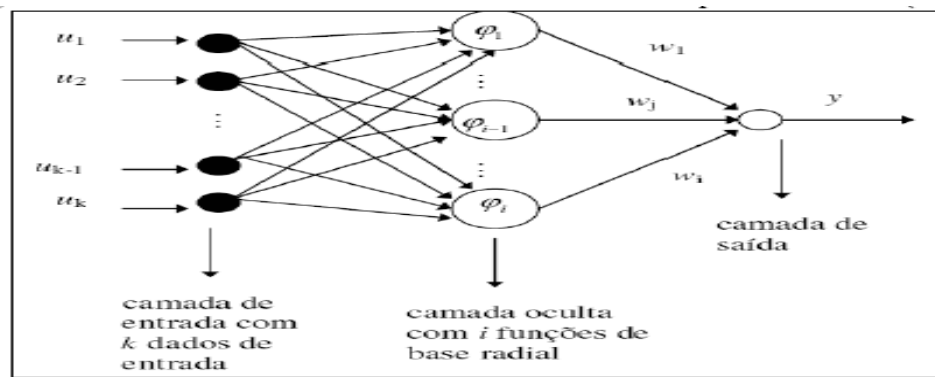


Fig. 8: Estrutura básica de uma rede neuronal RBF. Adaptado [26].

- **Modelos de referência**

No campo da previsão podem ser definidos alguns modelos simples que servem como uma referência ao avaliar a desempenho dos modelos desenvolvidos. Um dos modelos de referência mais utilizados é o modelo de persistência. Segundo [7] é um método de previsão *naive* onde o valor previsto de produção é o mesmo do último valor medido. Assim, este modelo pode ser definido pela seguinte equação:

$$\hat{P}_{t+k/t} = P_t \quad (3.3)$$

Onde $\hat{P}_{t+h/t}$ é a previsão para o horizonte de tempo $t + h$, calculado no instante t e P_t é a medição da potência de energia registrada no parque ou região em estudo.

Este modelo de previsão só tem um bom desempenho para uma previsão no horizonte muito pequeno, não mais do que 2 a 3 horas, embora, em previsão regional possa dar bons resultados para horizontes um pouco maiores, até 4 ou 6 horas.

O modelo de persistência é um caso particular do modelo de médias móveis nas últimas n observações, com $n = 1$

$$\hat{P}_{t+k/t} = \frac{1}{n} \sum_{kt=0}^{n-1} P_{t-kt} \quad (3.4),$$

sendo P_{t-kt} a medida de potência no instante $t-kt$.

Quando a ordem do modelo de média móvel tende para o infinito, então obtém-se a média geral de toda o histórico dos dados registados no parque ou região, também chamado de média climática:

$$\hat{P}_{t+k/t} = \bar{P}_t \quad (3.5),$$

onde \bar{P}_t é a média de todos os valores da potencia produzida até o instante t .

Também se pode encontrar modelo de previsão de potência eólica que pode ser considerado como referência e que resulta da combinação do modelo anterior (*persistência*), com as médias das potências registadas. Existe ainda um peso associado que é uma função de correlação, a_k entre P_t e P_{t+k} que é descrito segundo a equação (3.6):

$$\hat{P}_{t+k/t} = a_k * P_t + (1 - a_k) * \bar{P}_t \quad (3.6),$$

a_k é o fator de ponderação que deve ser estimado, com valores entre 0 e 1 dependendo do horizonte de previsão, e vai tomar valores próximos da unidade para previsão de horizontes de curto prazo e perto de zero para previsão de horizontes de longo prazo [10].

3.5. Modelos Híbridos

O modelo híbrido é um modelo que associa os modelos estatísticos das séries temporais com os modelos físicos.

Sendo um modelo híbrido, podem existir combinações de diferentes tipos:

- a) Combinação dos métodos físicos e estatísticos;
- b) Combinação dos modelos para curto prazo e médio prazo.

Qualquer modelo de previsão da energia eólica só poderá obter resultados satisfatórios se, à partida, obtiver melhores resultados, ou seja, menores erros na previsão, que os métodos considerados como sendo métodos de referência. O erro que relaciona os resultados previstos com a sua incerteza aumenta quando o prazo de previsão é alargado [13]. A Fig. 9 que se segue ilustra a representação de um modelo híbrido:

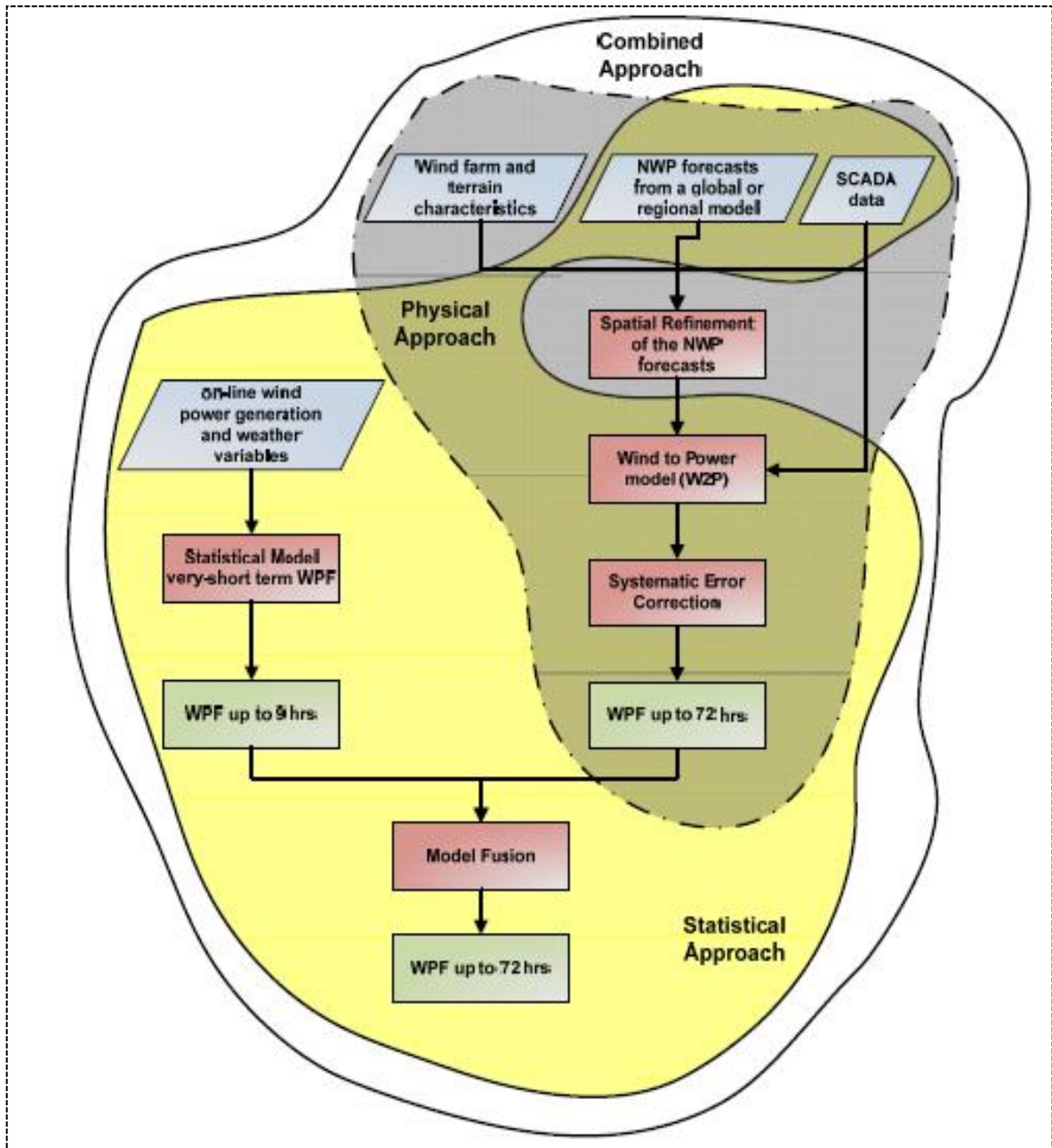


Fig. 9: Combinação dos diferentes modelos de previsão da energia eólica [13].

4. Previsão da Energia Eólica

4.1. Relação entre a Velocidade do Vento e a Potência

A figura que se segue representa o gráfico da curva de potência, relação entre a velocidade do vento v e a potência gerada P .

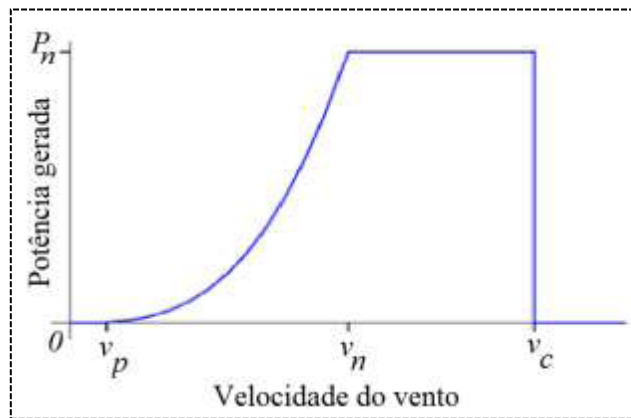


Fig. 10: Curva de potência típica de um aerogerador [16].

Conforme indicado pela curva de potência da Fig. 10, a conversão de energia eólica em energia elétrica dá-se a partir de uma velocidade mínima denominada velocidade de arranque V_p (*cut-in speed*). A potência da turbina é limitada ao valor nominal P_n , à velocidade nominal V_n (*rated speed*). Para velocidades entre V_p e V_n a potência elétrica P gerada pela turbina eólica cresce com o cubo da velocidade do vento v transversal à área A varrida pelo seu rotor dada pela equação (4.1) segundo [16]:

$$P = 0,5 * C_p * \eta_m * \eta_g * \rho * A * v^3 \quad (4.1)$$

Onde η_m é o rendimento mecânico da caixa de transmissão, η_g é o rendimento do gerador, ρ é a massa volúmica do ar e C_p é o coeficiente de potência da turbina, que expressa a parcela (%) da potência disponível do vento que é extraída pela turbina.

Ao modelar a curva de potência de um parque eólico, busca-se modelar as relações entre as variáveis de entrada, como por exemplo a velocidade do vento incidente nos aerogeradores do parque, e a variável de saída ou variável de resposta neste caso que é a eletricidade produzida por cada máquina ou parque eólico. Geralmente, a modelação da curva de potência é baseada na coleção de pares de dados de velocidade de vento e energia obtida e a obtenção de uma dispersão de pontos, como mostra a figura que se segue:

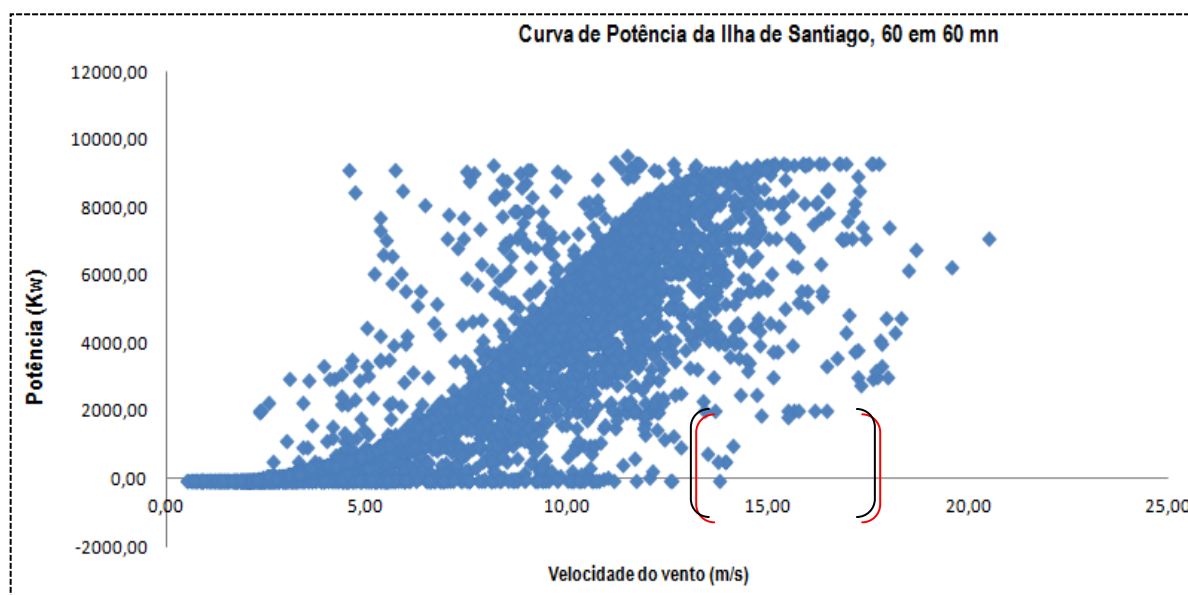


Fig. 11: Curva de potência que representa a produção em função da velocidade do vento do parque eólico da ilha de Santiago.

A curva de potência para os aerogeradores do parque em estudo (parque da ilha de Santiago) foi obtida da relação entre a média da velocidade do vento no intervalo de tempo de 60 minutos e a média horária da produção real de energia no mesmo intervalo de tempo. O coeficiente de correlação obtido foi de 0,813, o que nos permite afirmar que as variáveis estão fortemente correlacionadas. Por vezes, a velocidade do vento é superior à *cut-in speed*, (conferir as Fig. 10 e Fig. 11) mas a produção dos aerogeradores não corresponde à produção que seria esperada. Na figura supracitada (Fig. 11) são visíveis as situações de

indisponibilidade das máquinas, onde por vezes a sua produção é nula ou muito afastada do que deveria produzir segundo a curva de potência.

Também são visíveis as situações em que a velocidade do vento atinge o ponto *cut-out speed* e o decréscimo da potência produzida até zero.

Ainda da análise da curva de potência total produzida pelo parque constata-se situações em que apesar de a velocidade do vento medida apresentar um valor elevado, a produção total encontra-se muito afastada da potência nominal, isto porque poderão existir máquinas indisponíveis ou situações de velocidades extremas em que por razões de segurança é necessário desligar os aerogeradores.

A equação (4.1) indica que para prever a potência eólica gerada pode-se começar pela previsão da velocidade do vento. O modelo de persistência referenciado no capítulo 3.4 é o mais simples para prever a velocidade do vento e consiste em estimar a velocidade no instante $t+x$ como sendo a mesma do instante t . Naturalmente, o desempenho deste modelo decresce rapidamente à medida que o horizonte de previsão é aumentado. Em virtude da deficiência do modelo de persistência vários modelos para previsão de curto prazo da velocidade do vento têm sido desenvolvidos.

4.2. Previsão da Energia Eólica, Modelos Estatísticos

A fim de melhor potencializar o parque foram utilizados os dados de todos os aerogeradores (no total de 11) nas previsões. Para o efeito há diversos modelos possíveis de serem utilizados, porém, nem todos são apropriados para a nossa série temporal, pois resultariam em erros que poderiam comprometer todo o sistema de produção da energia do parque. Neste sentido serão apresentados alguns modelos os quais sejam:

- (i) *Holt Winters*, suavização exponencial (*exponential smoothing*);
- (ii) os modelos de metodologia de *Box-Jenkins* ou modelos *ARIMA* ;
- (iii) os modelos de redes neuronais artificiais.

A ideia não é colocar os modelos como competidores entre si, mas sim verificar a sua complementaridade, dado que podem ser muitas vezes usados conjuntamente ou alternativamente para a modelação de séries temporais. As descrições serão apresentadas tendo em consideração os aspetos probabilísticos, matemáticos e estatísticos utilizados para modelação e análise das séries temporais e a importância que desempenham na criação de estruturas matemáticas para representar o passado e prever o comportamento do futuro.

Os métodos propostos em (i) e (ii) não necessitam de quaisquer informações sobre as características de cada um dos aerogeradores, do terreno em que estão localizados, da potência das máquinas, da direção do vento ou de outros parâmetros meteorológicos do sistema operacional ao nível do parque em tempo real. A única informação necessária para a operação dos algoritmos de predições é ter o conjunto de n dados armazenados, dados históricos, contendo os valores das potências observadas em cada uma das máquinas instaladas. Para o modelo proposto em (iii) as variáveis de entrada do modelo são os valores previstos das condições meteorológicas de um modelo *Numerical Weather Prediction* (NWP) meso-escala (como por exemplo a velocidade do vento) e de saída os valores de produção (potência).

No trabalho serão utilizados três algoritmos de previsão implementados no pacote de previsão em R. O primeiro é baseado em modelos de alisamento (suavização) exponencial, o segundo o algoritmo para a previsão com modelos *ARIMA* e o terceiro redes neurais artificiais.

4.2.1. Medidas Estatísticas Padrões (ME, MAE, MSE, RMSE e MAPE)

Nesta secção é descrita as métricas frequentemente utilizadas na avaliação de modelos de regressão. Todas estas métricas são baseadas na diferença entre o valor previsto para o instante de tempo t e o valor observado nesse instante de tempo [20].

$$e_t = y_t - F_t \quad (4.2)$$

Para os casos onde apenas um valor de previsão é gerado chama-se de previsão de um passo à frente. O erro também recebe o nome de erro da previsão de um passo à frente.

Para os casos onde existe previsão para n períodos à frente e também n observações à frente poderão ser calculados n valores de erro, um para cada período. Para ambos os casos as seguintes equações de estatística padrão podem ser definidas:

Erro Médio (ME - *Mean Error*) dada pela equação (4.3), Erro Absoluto Médio (MAE = *Mean Absolute Error*), equação (4.4), Erro Quadrático Médio (MSE – *Mean Square Error*) dada pela equação (4.5), raiz quadrada do erro quadrático médio (RMSE- *Root Mean Squared Error*) equação (4.6) e Erro Médio Absoluto Percentual, (MAPE *Mean Absolute Percentage Error*) equação (4.9).

$$\text{Mean Error, ME} = \frac{1}{n} \sum_{t=1}^n e_t \quad (4.3)$$

$$\text{Mean Absolute Error, MAE} = \frac{1}{n} \sum_{t=1}^n |e_t| \quad (4.4)$$

$$\text{Mean Square Error, MSE} = \frac{1}{n} \sum_{t=1}^n (e_t)^2 \quad (4.5)$$

$$\text{Root Mean Squared Error, RMSE} = \sqrt{\text{MSE}} \quad (4.6)$$

A medida de precisão dessas estatísticas depende da escala dos dados utilizados; desta forma não permite a comparação entre séries temporais de diferentes áreas ou diferentes períodos.

Para comparar a precisão entre duas diferentes séries temporais ou intervalos de tempo é necessário usar uma medida do erro relativo ou percentual, em que são definidas da seguinte forma:

$$\text{Erro percentual, Percentage Error, PE}_t = [(y_t - F_t) / y_t] * 100 \quad (4.7)$$

$$\text{Erro Médio Percentual, Mean Percentage Error, MPE} = \frac{1}{n} \sum_{t=1}^n PE_t \quad (4.8)$$

Erro Médio Percentual Absoluto, *Mean Absolute Percentage Error*, MAPE

$$= \frac{1}{n} \sum_{t=1}^n |PE_t| \quad (4.9)$$

4.2.2. Modelo de Suavização Exponencial de Holt-Winters (ES)

Os modelos de suavização exponencial (*exponential smoothing*) foram proposto em 1950 por (Brown 1959, 1957 e Holt Winters 1960) e têm motivados alguns dos métodos de previsão. As previsões produzidas usam métodos de suavização exponencial das médias ponderadas das observações passadas da série temporal, sendo que os pesos decaem exponencialmente da mais recente a mais distante observação. Em outras palavras, a mais recente observação tem maior peso associado a uma taxa constante dada por:

$$a_j = \alpha(1-\alpha)^j, j=0,1,\dots; \text{ com } 0 < \alpha < 1 \quad (4.10)$$

Existem dois modelos de suavização exponencial utilizados no estudo das séries temporais: os modelos exponenciais de *Holt* e os modelos de suavização exponencial de Holt Winters. Uma vez que os modelos exponenciais de *Holt* são utilizados satisfatoriamente em séries temporais que apresentam tendência de crescimento linear (série em estudo tem comportamento global ascendente ou descendente) não será abordado nesse trabalho.

Os modelos de *Holt Winters* são modelos adequados para séries que apresentam tendências e sazonalidade. Tais modelos dividem-se em aditivo (SEHWA) e multiplicativo (SEHWM): no modelo aditivo parte-se do pressuposto que a amplitude da variação sazonal permanece constante ao longo do tempo e no modelo multiplicativo que a amplitude da variação sazonal varia ao longo do tempo, modelo cuja série em estudo mais se identifica (conferir. Fig. 13 da série).

As equações do modelo multiplicativo [17], equações (4.11), (4.12), (4.13) e (4.14) correspondem nível da série, tendência, sazonalidade e a previsão a ser realizada em K períodos a frente respetivamente:

$$L_t = \alpha \frac{Y_t}{St-s} + (1-\alpha)(L_{t-1} + T t-1) \quad (4.11)$$

$$T_t = \beta (L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (4.12)$$

$$St = \gamma \frac{Z_t}{L_t} + (1-\gamma)St-s \quad (4.13)$$

$$\hat{Y}_{t-k} = (L_t + K T_t) St-s-k \quad (4.14)$$

Análise estatística das medições da potência observada

A base de dados analisada foi fornecida pela empresa cabo-verdiana Cabeólica. Esta base contém registos das potências tomadas a cada 10 minutos. Os dados referem-se ao ano de 2012. A seguir, na tabela 4, são apresentadas algumas estatísticas descritivas da potência a cada 60 minutos:

Tabela 4: Tabela da estatística descritiva da Potência.

| | |
|-----------------------|-------------|
| Número de observações | 8436 |
| Média kW | 3340,9 |
| Mediana m/s | 2938,5 |
| Desvio-padrão | 2617,062 |
| Variância | 6849012,031 |
| Mínima kW | -32,0 |
| Máxima kW | 9720,0 |
| 1º Quartil | 958,8 |
| 3º Quartil | 5511,0 |

Na figura que se segue são apresentadas as distribuições das frequências das potências observadas a cada 60 minutos. As distribuições de frequências revelam que na maior parte do tempo o valor da potência não ultrapassa 8000 kW. O histograma também indica que a potência não segue uma distribuição Normal (estatística de *Kolmogorov-Smirnov*).

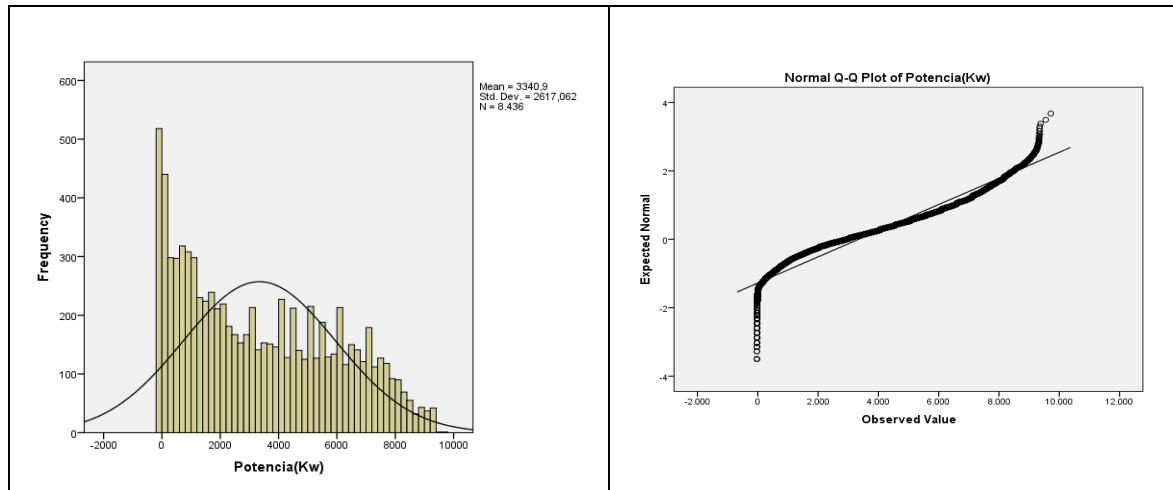


Fig. 12: Distribuição da frequência horária da potência do parque eólico da ilha de Santiago.

Para verificar o ajustamento da distribuição à normal (em que a hipótese nula, H_0 , nos diz que a variável em estudo segue distribuição normal, versus a hipótese alternativa, H_1 , a variável não segue uma distribuição normal, cuja regra de decisão é se $p\text{-value} < \alpha$, com $\alpha=0,05$, então rejeita-se H_0) aplicou-se o teste *K-S*. Na Fig.12 pode ainda verificar-se que os dados observados não seguem uma distribuição Normal ($p\text{-value} = 0,00$ em teste à normalidade: teste de *Kolmogorov-Smirnov*) como se pode constatar no *output* da tabela que se segue obtido através do *software IBM SPSS Statistics 21*:

Tabela 5: Teste de normalidade da variável potência.

| One-Sample Kolmogorov-Smirnov Test | | |
|--|----------------|----------|
| Power (kW) | | |
| N | | 8436 |
| Normal Parameters ^{a,b} | Mean | 3340,90 |
| | Std. Deviation | 2617,062 |
| Most Extreme Differences | Absolute | 0,102 |
| | Positive | 0,102 |
| | Negative | -0,099 |
| Kolmogorov-Smirnov Z | | 9,334 |
| Asymp. Sig. (2-tailed) | | 0,000 |
| ^{a.} Test distribution is Normal. | | |
| ^{b.} Calculated from data. | | |

Modelação da série temporal e implementação do algoritmo da previsão

Conforme foi anteriormente afirmado foi utilizado o modelo de suavização exponencial utilizando como critério de aferição das previsões, medidas dependentes das escalas que minimizem os erros.

O gráfico da figura que se segue, Fig. 13, ilustra o comportamento da série que representa o valor da potência de 60 em 60 minutos no parque eólico da ilha de Santiago. Os dados representam um total de 8.436 (oito mil, quatrocentos e trinta e seis) observações.

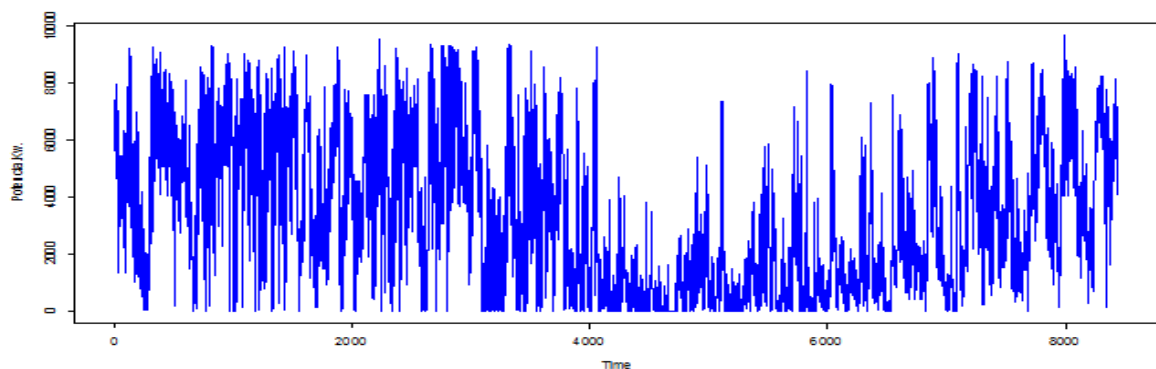


Fig. 13: Série da potência produzida de janeiro a dezembro de 2012. Os resultados são agregados à hora. Fonte: Elaborado pelo autor (2013) com base no banco de dados da empresa cabo-verdiana Cabeólica.

Pela observação do gráfico pode-se constatar que a série é não estacionária, apresentando padrões de comportamento diferenciados em diversos períodos, isto é, são encontrados períodos com menores variâncias e períodos com maiores variâncias. Para saber o ajustamento do modelo à série foi utilizado o pacote estatístico R versão 3.0.1. A modelação foi feita com base no algoritmo de suavização exponencial através do *ets* () do *package forecast* do Software R.

A tabela que se segue apresenta o resultado do valor do parâmetro estimado (potência) à série temporal com os seus 8.436 valores observados e com a previsão dos seus 24 valores futuros, utilizando o modelo *Holt Winters*, sendo o ME do ajustamento (MEa), o MAE do ajustamento (MAEa), o MAPE do ajustamento (MAPEa), ME da previsão (MEp), MAE da previsão (MAEp), e o MAPE da previsão (MAPEp):

Tabela 6: Valores dos parâmetros estimados utilizando o modelo *Holt Winters*, Alisamento Exponencial.

| Modelo | Parâm. | MEa | MAEa | MAPEa | MEp | MAEp | MAPEp | RMSEp |
|------------------------|---|---------------------|------|--------|-------|------|-------|-------|
| Alisamento Exponencial | α : 0,99 β : γ : 0,37 | 2,4e- ⁰⁵ | 0,22 | 177,27 | -0,30 | 0,46 | 76,20 | 0,54 |

Sendo, α : constante de ajuste do nível: β : constante de ajuste da tendência e γ : constante de ajuste da sazonalidade, os valores encontrados foram $\alpha = 0,99$ e $\gamma = 0,37$. Uma observação interessante é que como não se tem o valor para β isso poderá nos levar a afirmar que a série não apresenta tendência ou se tiver é muito fraca.

Observa-se que com a modelação da série encontra-se um valor muito baixo do erro médio (MEp), erro de previsão aproximadamente igual a -0,30, o que já não acontece com MAEp e MAPEp considerados relativamente altos, 0,46 e 76,20 respetivamente. Relativamente ao critério de informação do *Akaike*, foi obtido o seguinte valor: $AICc = 59077,95$.

As previsões foram feitas 24 horas à frente. Segue, Fig. 14, o gráfico das previsões com os respectivos dados da série:

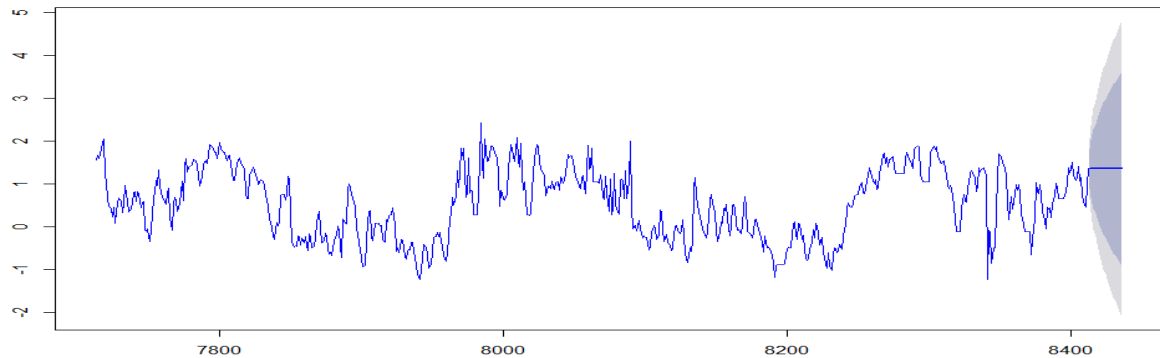


Fig. 14: Gráfico da potência observada e da potência prevista nas próximas 24 horas utilizando o modelo *Holt Winters*.

4.2.3. Modelo ARIMA

Este capítulo elucida a modelação da série, utilizando a função *auto-arma* () no R, para um modelo de séries temporais univariada, ou seja, que possui apenas uma variável descrevendo o seu comportamento em termos de seus valores passados. Foram experimentados vários modelos para ajustar o modelo *ARIMA*, mas chegou-se à conclusão que o modelo que aparentemente descreva a melhor sucessão cronológica da observação é a função *auto-arma* () descrito no capítulo 3.4.

Para a escolha do modelo com melhor desempenho na previsão da nossa série nas próximas 24 horas foram realizados testes com o *auto-arma* e com vários outros modelos *ARIMA* que adaptam a série a fim de comparar os resultados. Os resultados obtidos encontram-se ilustrados na tabela que se segue seguindo o quadro comparativo dos *AICc*:

Tabela 7: Alguns modelos adequados utilizados na modelação das séries.

| | Modelo | Valor do AICc |
|---------------------|--------------------------|---------------|
| <i>auto-arima</i> | ARIMA _(4,1,1) | 6414,36 |
| Modelo ₁ | ARIMA _(1,0,1) | 6931,62 |
| Modelo ₂ | ARIMA _(1,1,1) | 6521,87 |
| Modelo ₃ | ARIMA _(1,2,1) | 6965,70 |
| Modelo ₄ | ARIMA _(0,1,1) | 9010,28 |
| Modelo ₅ | ARIMA _(1,2,5) | 6465,32 |

Pela tabela 7 podemos constatar que a função *auto-arima* tem o *AICc* mais baixo de que todos os outros modelos escolhidos. O *AICc* do *auto-arima* (4,1,1) é igual a 6414,36, enquanto que para o modelo *ARIMA* o que apresenta o *AICc* mais baixo é o modelo *ARIMA*(1,2,5) com o *AICc* igual a 6465,32. Logo justifica a escolha da função *auto-arima* para modelação da série.

Modelação da série temporal e implementação do algoritmo da previsão

Dada à dimensão dos dados (oito mil quatrocentos e trinta e seis observações) construiu-se o gráfico da *FAC* com 400 *lags* permitindo assim ter uma melhor visibilidade do seu comportamento. Da observação da Fig. 15, constata-se que o gráfico da Autocorrelação (*FAC*) decresce até um certo instante e volta a crescer. Os pontos considerados máximos correspondem aos pontos dos ciclos diários da série.

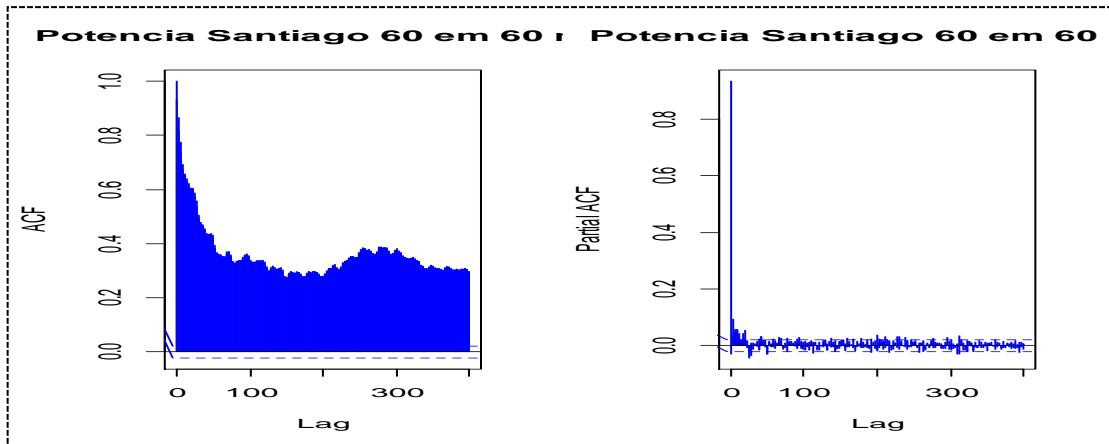


Fig. 15: Gráficos da ACF e PACF.

Os dados são claramente não-estacionários. A *FAC* corta na ordem do processo e a *FACP* decresce para zero na parte sazonal; e na componente não sazonal *FAC* e *FACP* não tem decaimento. Aplicando a função *auto.arima* () obteve-se o modelo que melhor se parece ajustar à série; *ARIMA* (4,1,1), ou seja um modelo descrito por quatro parâmetros autorregressivos a uma diferença e um parâmetro de média móvel.

Tem-se então o gráfico obtido pela utilização de um modelo *ARIMA* obtido pela função *auto-arima* para a previsão nas próximas 24 horas como mostra a Fig. 16.

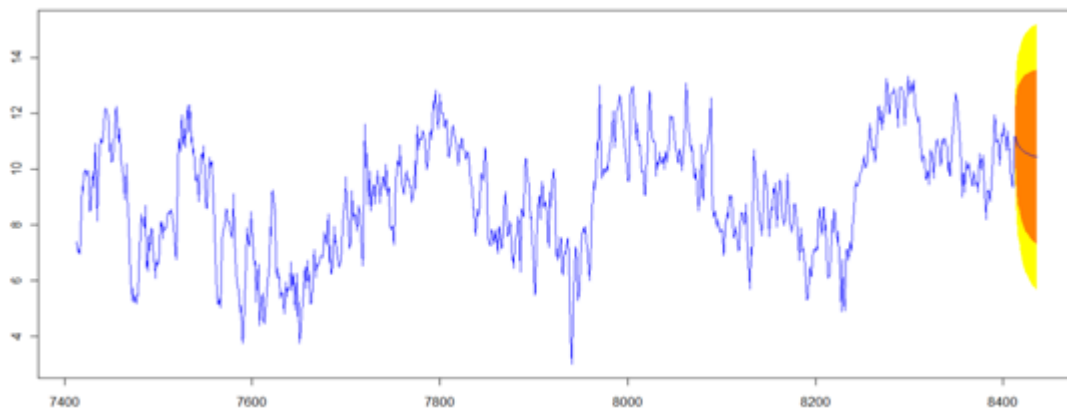


Fig. 16: Gráfico da potência observada e da potência prevista nas próximas 24 horas utilizando o modelo *ARIMA* (4, 1, 1).

Resultado da previsão da série temporal

Nesta seção é apresentada a previsão da série, considerando os modelos que apresentam os melhores resultados. Enquanto o *exponential smoothing* os modelos foram baseados em uma descrição da tendência e sazonalidade dos dados, os modelos *ARIMA* visam descrever a autocorrelação nos dados.

A tabela que se segue ilustra os resultados, das previsões, obtidos pelo método de alisamento exponencial e o modelo *ARIMA* nas próximas 24 horas:

Foram obtidos os seguintes erros: ME, MAE, MAPE e o RMSE das previsões:

Tabela 8: Comparação dos erros entre Holt- Winters e ARIMA.

| Modelos | | ME | MAE | MAPE | RMSE |
|------------------------|------------------------|--------|-------|--------|-------|
| Alisamento Exponencial | <i>H. Winters</i> | -0,301 | 0,463 | 76,202 | 0,541 |
| ARIMA | <i>ARIMA</i> $(4,1,1)$ | -0,005 | 0,373 | 51,351 | 0,448 |

Análise do Erro

Como já tinha sido referenciado no capítulo 4.2.1, para comparação do desempenho de modelos na previsão entre duas diferentes séries temporais ou intervalos de tempo será usada a medida do erro percentual, MAPE. Onde MAPE é o Erro Médio Absoluto Percentual, *Mean Absolute Percentage Error*.

De acordo com a tabela 8, dos erros das previsões, pode-se verificar que o Erro Médio Absoluto Percentual obtido pelo modelo *ARIMA* $(4,1,1)$ é inferior ao erro obtido pelo método de *Holt Winters*, sendo que para o *ARIMA* $(4,1,1)$ o valor do MAPE é igual a 51,4% e *Holt Winters* é igual a 76,2%. Logo conclui-se que o modelo *ARIMA* $(4,1,1)$ teve melhor desempenho que o *Holt Winters*.

4.2.4. Redes Neurais Artificiais – RNAs

4.2.4.1. Perceptron de Múltiplas Camadas (RNAs- MLP)

Neste capítulo do trabalho o objetivo é utilizar modelos de Redes Neurais Artificiais (RNAs) na previsão de potências voltadas à previsão da produção da energia eólica. Na modelação foram utilizados dois tipos de redes neuronais: *Multi Layer Perceptrons (MLP)* e *Radial Basis Function (RBF)*. A estatística utilizada para caraterizar os modelos é a análise do Erro Médio (ME), Erro Médio Absoluto (MAE), Erro Médio Absoluto percentual (MAPE) e raiz quadrada do erro quadrático médio (RMSE).

Mas segundo [28] é conveniente introduzir o erro de previsão normalizado, que consiste apenas na divisão do erro pela potência instalada (P_{inst}). É importante referir que uma normalização deste género conduz a valores de erro inferiores. No entanto, pode-se registar algumas falhas quando se analisa o erro, critério esse utilizado na modelação das RN.

Assim o erro pode ser definido da seguinte forma:

$$e_t = \frac{1}{P_{inst}} (y_t - F_t) \quad (4.15),$$

sendo P_{inst} : capacidade do parque (potência máxima instalada).

O 1º modelo de RNA apresentado neste estudo é *Multi Layer Perceptron (MLP)*. A modelação da rede foi feita seguindo o seguinte critério de validação cruzada:

- (i) Os dados foram divididos em dois conjuntos: conjunto de treino e conjunto de teste;
- (ii) Os conjuntos de treino foram divididos em: **conjunto treino 1**- meses de janeiro a setembro, **conjunto treino 2** - meses de abril a dezembro e **conjunto treino 3** - meses de janeiro a maio e de setembro a dezembro e para o conjunto de teste foram: **conjunto teste 1**- meses de outubro a dezembro, **conjunto teste 2** - meses de janeiro a março e **conjunto teste 3** - meses de junho a agosto.

Na base do sistema de previsão estão as previsões meteorológicas obtidas a partir do modelo NWP e o histórico de produção da potência do parque.

As variáveis de entrada utilizadas são:

- Velocidade do vento;
- Velocidade do vento normalizada;
- Direção do vento
- Componente seno da direção do vento;
- Componente cosseno da direção do vento;
- Temperatura;
- Massa volúmica do ar;
- Pressão atmosférica.

De modo a poder ter uma visualização melhor das informações na tabela que apresenta os valores dos erros obtidos foram consideradas as seguintes designações para as variáveis em estudado:

Tabela 9: Descrição das siglas.

| Ordem | Designação | Sigla |
|-------|---|---|
| 1 | Velocidade e direção do vento | VDV |
| 2 | Velocidade do vento, direção do vento e temperatura | VDVT |
| 3 | Velocidade do vento, direção do vento, temperatura e pressão | VDVTP |
| 4 | Velocidade do vento, velocidade do vento normalizada, direção do vento e temperatura e pressão | VV _N DVTP |
| 5 | Velocidade do vento, velocidade do vento normalizada, direção do vento, direção seno, temperatura e pressão | VV _N DVD _S TP |
| 6 | Velocidade do vento, velocidade do vento normalizada, direção do vento, direção seno, direção cosseno, temperatura e pressão | VV _N DVD _S D _C TP |
| 7 | Velocidade do vento, velocidade do vento normalizada, direção do vento, direção seno, direção cosseno, temperatura, pressão e massa do ar | VV _N DVD _S D _C TPM |

A configuração das *RNAs*, *Multi Layer Perceptrons* (MLP) foi feita utilizando três camadas, com apenas uma camada escondida, com uma taxa de aprendizagem de 0,01 e com 500 iterações. O número de neurónios na camada escondida varia de acordo com o número de variáveis de entrada ($2n + 1$) em que o n representa o número de variáveis. Nos dados de entrada, optou-se por acrescentar apenas uma variável em cada simulação, evitando ajustar mais do que uma variável em simultâneo e assim possuir um número infindável de combinações.

Os resultados que se seguem, tabela 10, dizem respeito à modelação da série temporal utilizando o modelo *RNAs*, *Multi Layer Perceptrons*. O processo de modelação foi feito utilizando o Software R versão 3.0.1 (2013-05-16). Os dados do conjunto de treino foram referentes aos meses de janeiro a setembro e os de conjunto de teste dos meses de outubro a dezembro. Assim, a tabela que se segue apresenta os resultados da variação dos *Inputs* utilizados na modelação de série e os erros que auxiliaram na escolha do modelo que melhor descreve o processo gerador da série temporal.

Tabela 10: Modelação da série temporal utilizando Multi Layer Perceptrons.

| Input | MLP | NME | NMAE | NMAPE | NRMSE | Correlação |
|--|----------|-------|-------|--------|-------|------------|
| VDV | (2,5,1) | 0,176 | 0,426 | 42,640 | 0,509 | 0,830 |
| VDVT | (3,7,1) | 0,054 | 0,400 | 40,065 | 0,503 | 0,845 |
| VDVTP | (4,9,1) | 0,059 | 0,415 | 41,598 | 0,497 | 0,819 |
| VV _N DVTP | (5,11,1) | 0,070 | 0,453 | 45,388 | 0,597 | 0,815 |
| VV _N DVD _S TP | (6,13,1) | 0,048 | 0,392 | 39,292 | 0,487 | 0,830 |
| VV _N DVD _S DC _{TP} | (7,15,1) | 0,074 | 0,423 | 42,348 | 0,267 | 0,816 |
| VV _N DVD _S DC _{TPM} | (8,17,1) | 0,106 | 0,455 | 45,545 | 0,567 | 0,797 |

Pela tabela supra verifica-se que o modelo de *RNAs*, que melhor descreveu o processo gerador da série temporal foi o *MLP* (6,13,1) ou seja a rede com seis neurónios na camada de entra, treze neurónios na camada intermédia e um neurónio na camada de saída, com os seguintes *inputs*: velocidade do vento, velocidade do vento normalizada, direção do vento, direção seno, temperatura e pressão. Ao iniciar o treino da rede com as variáveis: velocidade do vento, velocidade do vento normalizada, direção do vento, temperatura e

pressão, o acréscimo da variável, direção do vento seno, proporcionou melhoria preditiva da rede.

O gráfico da Fig. 17, ilustra os resultados dos erros obtidos no processo de teste para todos os modelos de *RNAs-MLP*.

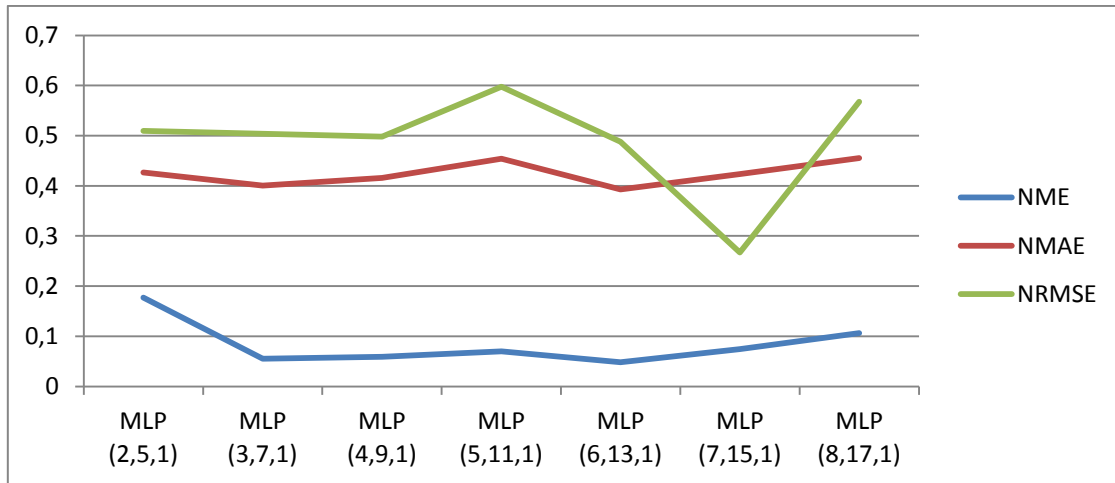


Fig. 17: Gráfico de teste de todas as redes MLP (outubro a dezembro de 2012).

Segue o gráfico que ilustra a desempenho dos modelos de *RNAs MLP*. Como critério foi utilizado o erro médio absoluto percentual normalizado e a rede *MLP (6,13,1)* obteve um erro de 39,3 %, sendo o menor obtido e a correlação entre os dados observados e os previstos é de 0,830.

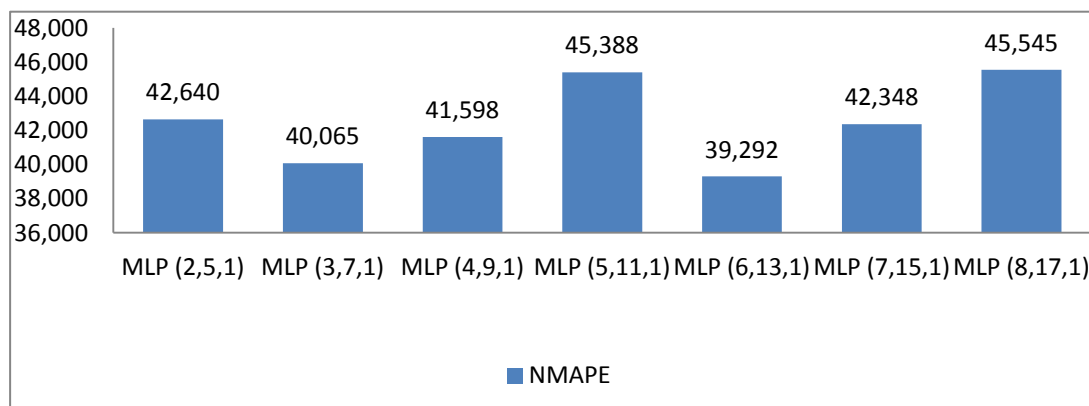


Fig. 18: Gráfico do desempenho dos modelos de *RNAs-MLP* (outubro a dezembro de 2012).

Desta forma, a Fig. 19 ilustra a arquitetura da rede *MLP* (6,13,1) com o melhor desempenho e a composição da mesma.

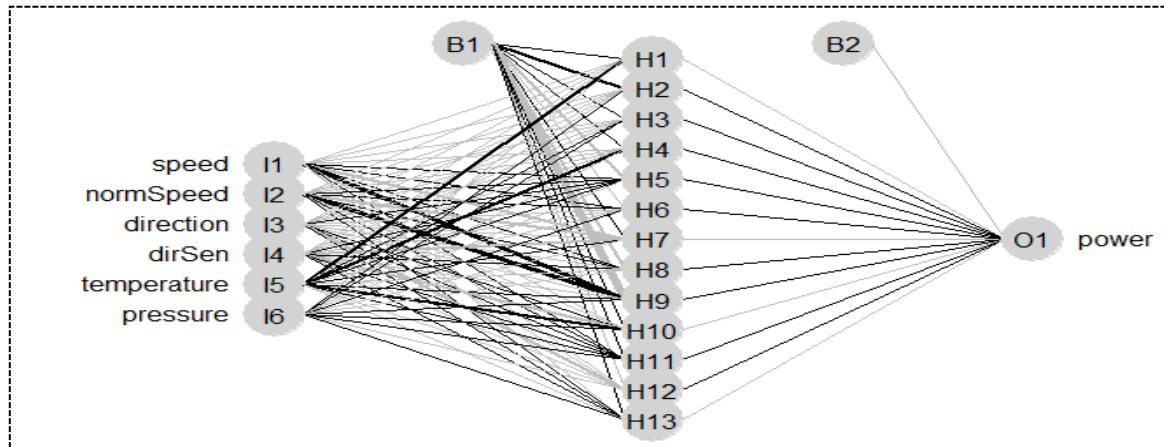


Fig. 19: Rede MLP (6,13,1) utilizada na modelação da série.

Os resultados que se seguem dizem respeito à modelação da série temporal utilizando o modelo de *RNAs*, *Multi Layer Perceptrons*, em que o conjunto de treino contem os dados dos meses de abril a dezembro e o conjunto de teste os meses de janeiro a março. Assim a tabela 11 apresenta os resultados da variação dos *Inputs* utilizados na modelação de série e os erros que auxiliaram na escolha do modelo que melhor descreve o processo gerador da série temporal.

Tabela 11: Modelação da série temporal utilizando Multi Layer Perceptrons.

| Input | MLP | NME | NMAE | NMAPE | NRMSE | Correlação |
|--|----------|-------|-------|--------|-------|------------|
| VDV | (2,5,1) | 0,236 | 0,540 | 54,040 | 0,643 | 0,756 |
| VDVT | (3,7,1) | 0,312 | 0,554 | 55,464 | 0,660 | 0,751 |
| VDVTP | (4,9,1) | 0,281 | 0,552 | 55,229 | 0,649 | 0,743 |
| VV _N DVTP | (5,11,1) | 0,31 | 0,542 | 54,208 | 0,655 | 0,721 |
| VV _N DVD _S TP | (6,13,1) | 0,281 | 0,534 | 53,481 | 0,634 | 0,720 |
| VV _N DVD _S DC _{TP} | (7,15,1) | 0,356 | 0,584 | 58,457 | 0,677 | 0,732 |
| VV _N DVD _S DC _{TPM} | (8,17,1) | 0,347 | 0,583 | 58,317 | 0,678 | 0,740 |

Pela tabela 11 e pelo gráfico da Fig. 20 verifica-se que o modelo da *RNAs*, que melhor descreveu o processo gerador da série temporal foi o *MLP* (6,13,1) com treze neurónios na

camada interna e cujos *inputs* foram: velocidade do vento, velocidade do vento normalizada, direção do vento, direção seno, temperatura e pressão.

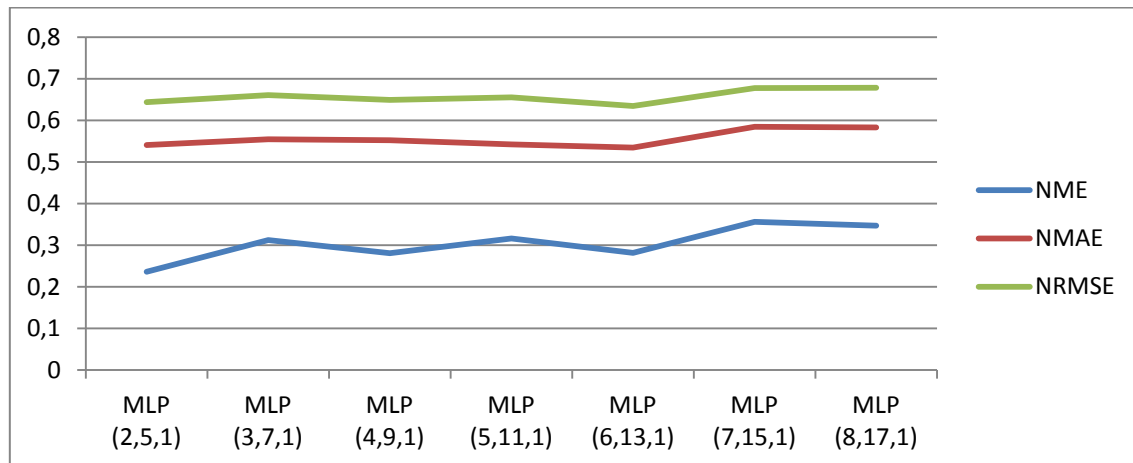


Fig. 20: Gráfico de teste de todas as redes MLP (janeiro a março de 2012).

Segue o gráfico que ilustra o desempenho dos modelos de *RNAs MLP*. Como critério foi utilizado o erro médio absoluto percentual e a rede *MLP (6,13,1)* obteve um erro de 53,5 %, sendo o menor obtido e a correlação entre os dados observados e os previstos é de 0,720, o que significa que os erros estão bem correlacionados, ou seja há uma correlação forte entre os dados observados e os previstos.

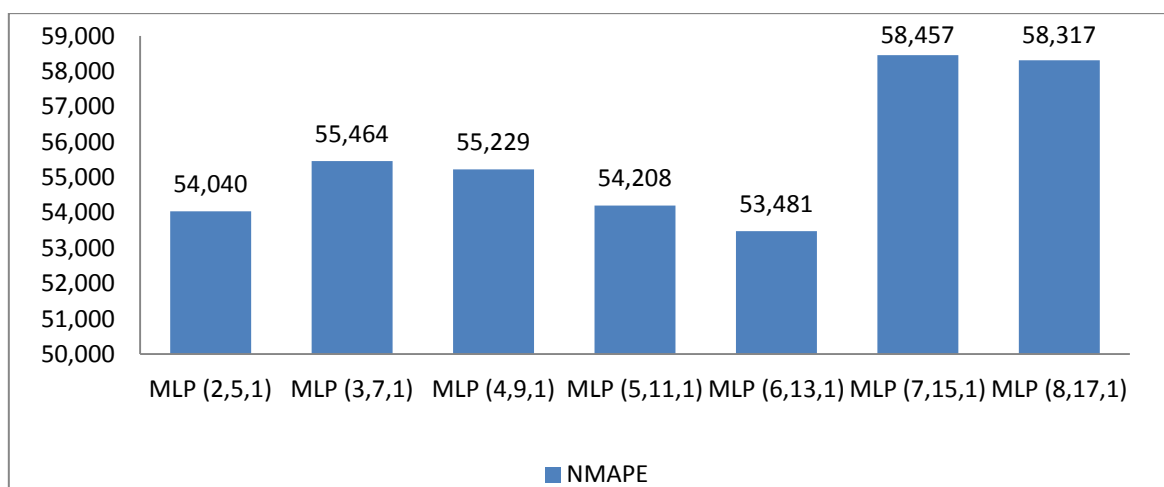


Fig. 21: Gráfico do desempenho dos modelos de *RNAs-MLP* (janeiro a março de 2012).

Observa-se que, com a modelação da série utilizando o intervalo de tempo para teste, meses de outubro a dezembro, os modelos não capturaram adequadamente o processo gerador da série. Verifica-se isto, pois, os valores dos erros, absoluto médio percentual, estão entre 53,5 % e 58,5 % e nenhum dos erros de previsão (erro médio, erro médio absoluto e raiz quadrada do erro quadrático médio) converge para o erro meta estipulado (0,01), pois as previsões das RNAs não atingiram valores aceitáveis de erros.

A Fig. 22 apresenta a rede *MLP* (6,13,1) ilustrando a composição da mesma, ou seja, uma rede com seis neurónios na camada de entrada, treze neurónios na camada intermédia e um neurónio na camada de saída.

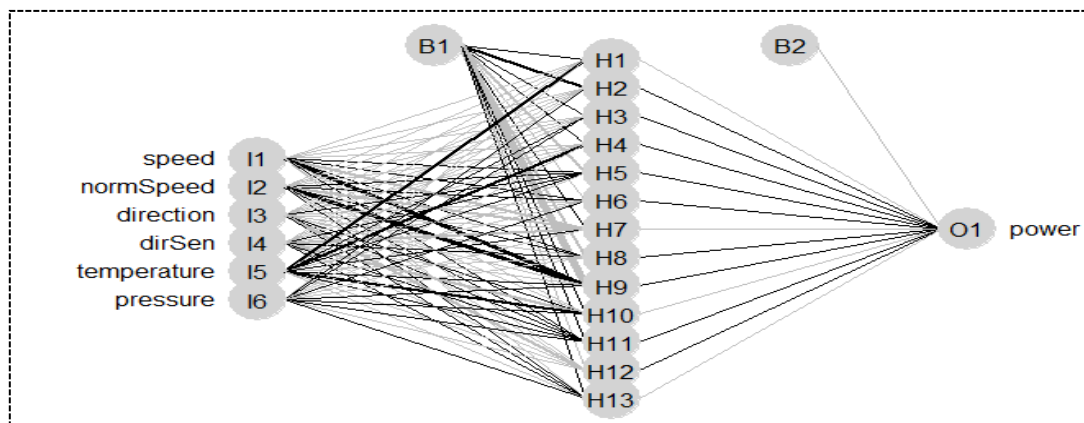


Fig. 22: Rede MLP (6,13,1) utilizada na modelação da série.

De seguida é apresentada a modelação da série temporal utilizando o modelo de *RNAs*, *Multi Layer Perceptrons* em que o conjunto de treino contém os dados referentes aos meses de janeiro a maio e de setembro a dezembro, conjunto de treino e os de conjunto de teste, meses de junho a agosto.

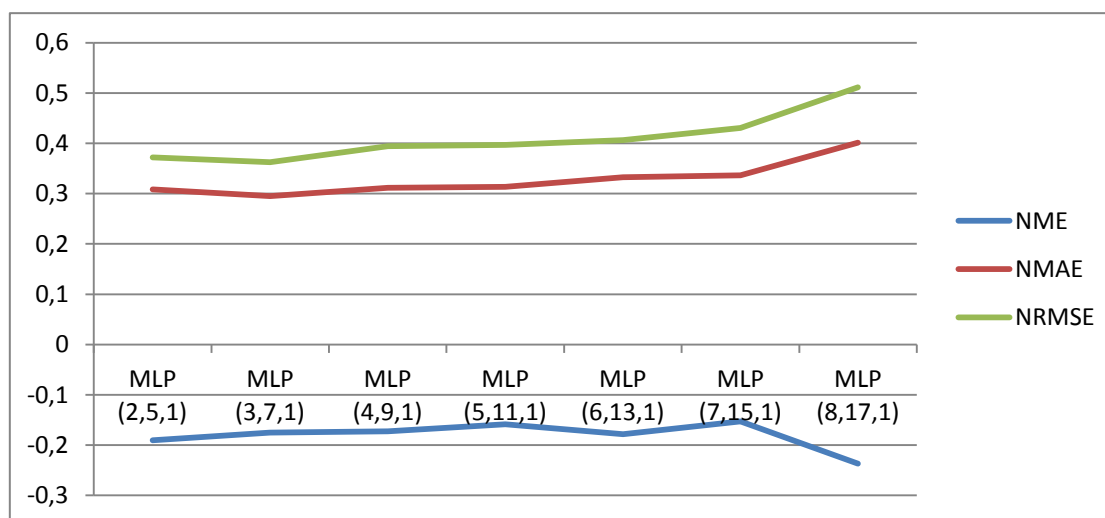
A tabela 12 apresenta os resultados da variação dos *Inputs* utilizados na modelação de série e os erros que auxiliaram na escolha do modelo que melhor descreve o processo gerador da série temporal.

Tabela 12: Modelação da série temporal utilizando Multi Layer Perceptrons.

| Input | MLP | NME | NMAE | NMAPE | NRMSE | Corel. |
|---|----------|--------|-------|--------|-------|--------|
| VDV | (2,5,1) | -0,190 | 0,308 | 30,817 | 0,372 | 0,702 |
| VDVT | (3,7,1) | -0,175 | 0,295 | 29,528 | 0,362 | 0,671 |
| VDVTP | (4,9,1) | -0,172 | 0,311 | 31,161 | 0,394 | 0,643 |
| VV _N DVTP | (5,11,1) | -0,158 | 0,313 | 31,343 | 0,396 | 0,616 |
| VV _N DVD _S TP | (6,13,1) | -0,178 | 0,332 | 33,281 | 0,406 | 0,639 |
| VV _N DVD _S D _C TP | (7,15,1) | -0,152 | 0,336 | 33,656 | 0,430 | 0,619 |
| VV _N DVD _S D _C TPM | (8,17,1) | -0,236 | 0,401 | 40,110 | 0,511 | 0,599 |

O modelo da *RNA*, que melhor descreveu o processo gerador da série temporal foi o *MLP* (3,7,1) com três neurónios na camada de entrada, sete neurónios na camada intermédia e um neurónio na camada de saída e cujos *Inputs* foram: velocidade do vento, direção do vento e temperatura. A não agregação ou agregação de quaisquer outras variáveis ao grupo levou a piores resultados dos modelos.

A figura infra ilustra os resultados dos erros obtidos no processo de teste para os sete modelos de *RNAs*, com os diferentes *Inputs*:

**Fig. 23: Gráfico do teste de todas as *RNs-MLP* (junho a agosto de 2012).**

O gráfico da figura que segue ilustra o desempenho dos modelos de *RNAs*. Como critério foi utilizado o erro médio absoluto percentual normalizado e a rede *MLP* (3,7,1) é a que obteve melhor resultado, com o *NMAPE* de 29,5%, sendo o menor obtido e a correlação entre os dados observados e os previstos é de 0,671, o que significa que os erros estão correlacionados.

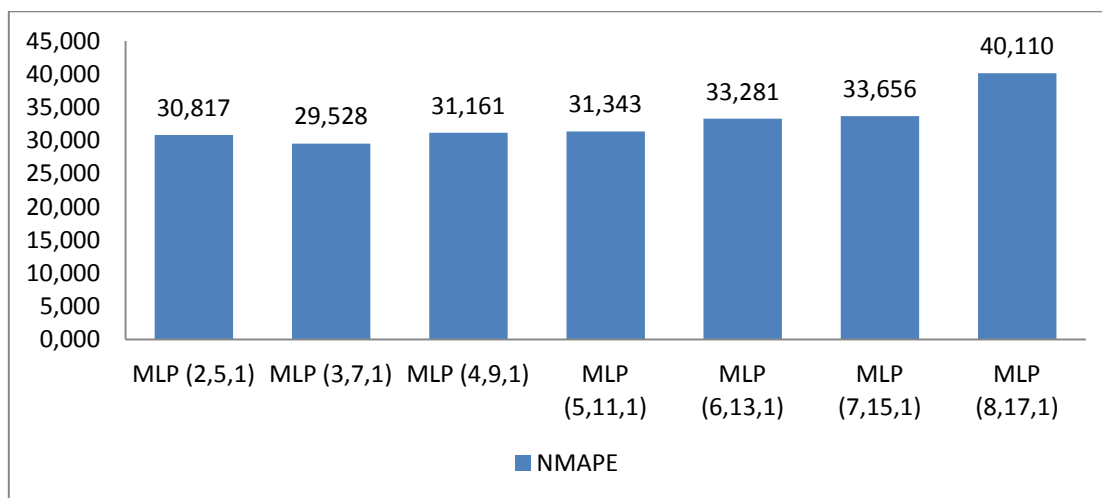


Fig. 24: Gráfico do desempenho dos modelos de *RNAs-MLP*, (junho a agosto).

Observa-se que, com a modelação da série utilizando o intervalo de tempo para teste, meses de junho a agosto, o modelo *MLP* (3,7,1) consegue capturar o processo gerador da série. Verifica-se isto, pois, o valor do erro, absoluto médio percentual normalizado é de 29,5%, valor considerado aceitável, mas nenhum dos erros de previsão (erro médio, erro médio absoluto e raiz quadrada do erro quadrático médio) converge para o erro meta estipulado (0,01).

A Fig. 25 apresenta a rede *MLP* (3,7,1) ilustrando a composição da mesma, ou seja, uma rede com três neurónios na camada de entrada, sete na camada intermédia e um na camada de saída.

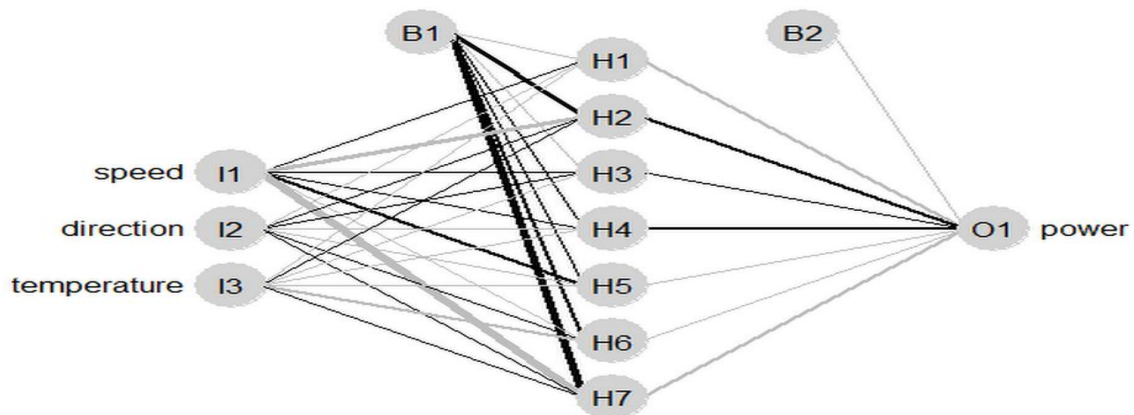


Fig. 25: Rede MLP (3,7,1) utilizada na modelação da série.

Também fez-se a modelação para o cálculo dos erros para as 24 redes *MLP* com todas as variáveis de entrada. Assim as redes foram configuradas seguindo os seguintes detalhes: nº de neurónios na camada de entrada igual a 3, nº de neurónios na camada escondida igual a 7 e nº de neurónios na camada de saída igual a 1.

Seguem os resultados das modelações das vinte e quatro redes *Multi Layer Perceptron* (*MLP*), em que cada uma foi feita no intervalo de tempo de uma hora, ou seja as vinte e quatro redes correspondem a vinte e quatro horas respetivamente, seguindo os seguintes pressupostos:

(i) os dados foram divididos em dois conjuntos: conjunto de treino correspondente a 70 % do total dos dados e os restantes 30% para conjunto de teste; os *Inputs* foram todas as variáveis utilizadas no estudo.

Tabela 13: Erros obtidos na modelação das 24 redes MLP.

| Rede | Input | MLP | NME | NMAE | NMAPE | NRMSE | Correl. |
|------|---------------|----------|--------|--------|--------|-------|---------|
| R 1 | VVn DVDSDcTPM | (8,17,1) | -0,084 | 0,677 | 67,742 | 0,884 | 0,646 |
| R 2 | VVn DVDSDcTPM | (8,17,1) | -0,058 | 0,660 | 66,061 | 0,806 | 0,705 |
| R 3 | VVn DVDSDcTPM | (8,17,1) | 0,139 | 0,622 | 62,223 | 0,823 | 0,648 |
| R 4 | VVn DVDSDcTPM | (8,17,1) | 0,056 | 0,646 | 64,639 | 0,866 | 0,535 |
| R 5 | VVn DVDSDcTPM | (8,17,1) | -0,049 | 0,6352 | 63,528 | 0,864 | 0,662 |
| R 6 | VVn DVDSDcTPM | (8,17,1) | 0,078 | 0,7428 | 74,288 | 0,944 | 0,511 |
| R 7 | VVn DVDSDcTPM | (8,17,1) | -0,102 | 0,764 | 76,493 | 0,946 | 0,542 |
| R 8 | VVn DVDSDcTPM | (8,17,1) | 0,125 | 0,604 | 60,488 | 0,783 | 0,679 |
| R 9 | VVn DVDSDcTPM | (8,17,1) | -0,023 | 0,666 | 66,655 | 0,829 | 0,590 |
| R 10 | VVn DVDSDcTPM | (8,17,1) | 0,014 | 0,636 | 63,698 | 0,827 | 0,658 |
| R 11 | VVn DVDSDcTPM | (8,17,1) | -0,138 | 0,700 | 70,069 | 0,881 | 0,582 |
| R 12 | VVn DVDSDcTPM | (8,17,1) | -0,041 | 0,694 | 69,412 | 0,882 | 0,589 |
| R 13 | VVn DVDSDcTPM | (8,17,1) | -0,113 | 0,594 | 59,490 | 0,749 | 0,610 |
| R 14 | VVn DVDSDcTPM | (8,17,1) | 0,143 | 0,587 | 58,786 | 0,827 | 0,710 |
| R 15 | VVn DVDSDcTPM | (8,17,1) | -0,179 | 0,739 | 73,920 | 0,987 | 0,586 |
| R 16 | VVn DVDSDcTPM | (8,17,1) | 0,171 | 0,674 | 67,489 | 0,944 | 0,705 |
| R 17 | VVn DVDSDcTPM | (8,17,1) | -0,063 | 0,676 | 67,682 | 0,875 | 0,499 |
| R 18 | VVn DVDSDcTPM | (8,17,1) | 0,055 | 0,605 | 60,504 | 0,787 | 0,482 |
| R 19 | VVn DVDSDcTPM | (8,17,1) | 0,023 | 0,639 | 63,980 | 0,858 | 0,549 |
| R 20 | VVn DVDSDcTPM | (8,17,1) | 0,159 | 0,662 | 66,296 | 0,895 | 0,604 |
| R 21 | VVn DVDSDcTPM | (8,17,1) | -0,015 | 0,676 | 67,621 | 0,869 | 0,664 |
| R 22 | VVn DVDSDcTPM | (8,17,1) | 0,013 | 0,670 | 67,000 | 0,866 | 0,647 |
| R 23 | VVn DVDSDcTPM | (8,17,1) | 0,117 | 0,661 | 66,191 | 0,854 | 0,576 |
| R 24 | VVn DVDSDcTPM | (8,17,1) | 0,099 | 0,630 | 63,067 | 0,795 | 0,525 |

Todas as vinte e quatro redes têm a mesma arquitetura como podemos visualizar na Fig. 26, *MLP (8,17,1)*, com oito neurónios na camada de entrada, dezassete neurónios na camada intermédia e um na camada de saída. Pois têm as mesmas variáveis utilizadas na

modelação, velocidade do vento, velocidade do vento normalizada, direção do vento, direção do vento seno, direção do vento cosseno, temperatura, pressão e massa do ar.

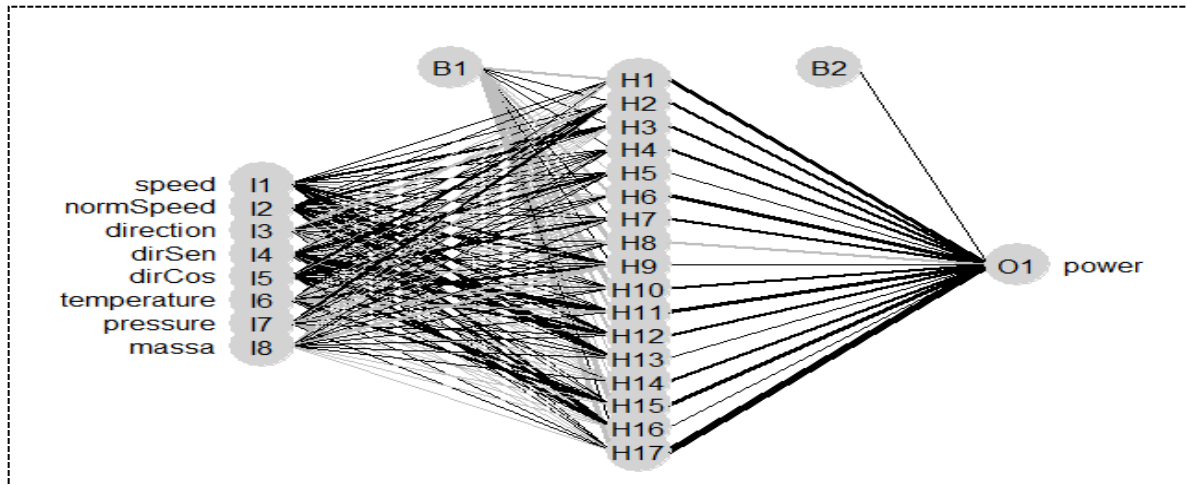


Fig. 26: Rede MLP (8,17,1) utilizada na modelação da série.

Podemos constatar que entre as redes modeladas, a que produz melhores resultados é a *R14*, rede modelada no tempo, $t = 14$ horas, como ilustram os gráficos das Fgs. 27 e 28. A rede com melhor desempenho apresenta o erro médio normalizado, NME igual a 14,4%, erro absoluto médio normalizado, NMAE igual a 58,8, raiz quadrada do erro quadrático médio normalizado, NRMSE igual a 82,8% e erro absoluto médio percentual normalizado igual a 58,8%.

O gráfico que se segue ilustra os erros obtidos na modelação das vinte e quatro redes:

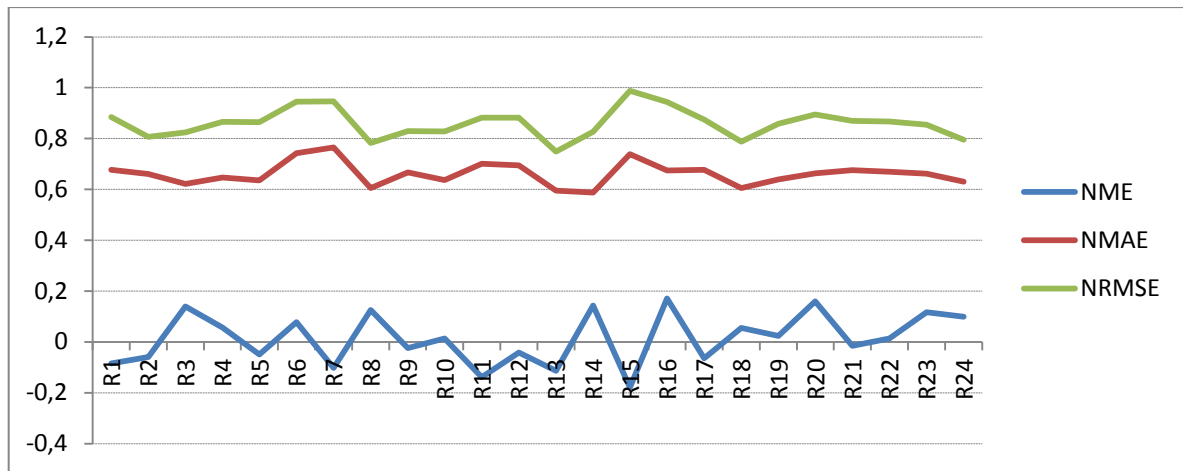


Fig. 27: Gráfico dos erros obtidos na previsão da potência pelas 24 redes MLP.

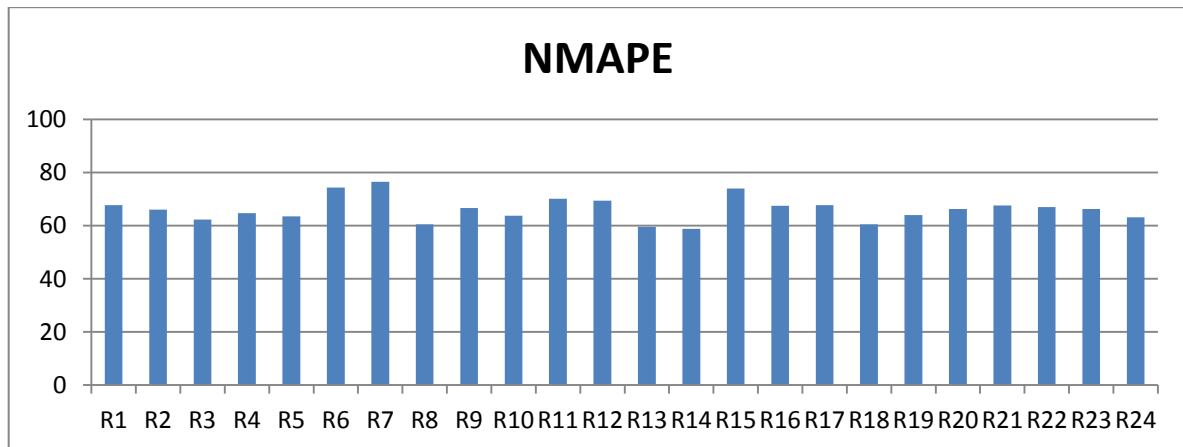


Fig. 28: Gráfico dos erros obtidos na previsão da potência pelas 24 redes MLP

Observa-se que, com a modelação da série utilizando vinte e quatro redes, os modelos não capturaram adequadamente o processo gerador da série. Verifica-se isto, pois, os valores dos erros absolutos médios percentuais estão entre 58,8% e 76,5%, valores que podem ser considerados altos. Mas pode-se observar que o erro médio da previsão atingiu o critério de convergência (0,01) para as 500 iterações para $t = 10$ horas e $t = 22$ horas.

Na Fig. 29 é apresentado o gráfico síntese dos resultados dos erros dos modelos das *RNAs MPL* com melhores desempenhos.

Através da análise dos erros obtidos, NMAPE, na previsão da “potência” pelas redes *MLP*, podemos constatar que de entre as redes modeladas, a que produz melhores resultados é a

rede *MLP* (3,7,1), com os parâmetros de entrada, velocidade e direção do vento e temperatura, ou seja os melhores resultados da previsão da potência encontram-se associados aos fenómenos meteorológicos, velocidade do vento, direção do vento e temperatura nos meses de janeiro a maio e de setembro a dezembro (conjunto de treino) e dos meses de junho a agosto (conjunto de teste), sendo erro médio absoluto percentual normalizado, NMAPE de 29,5%, como podemos verificar através do gráfico da figura que se segue:

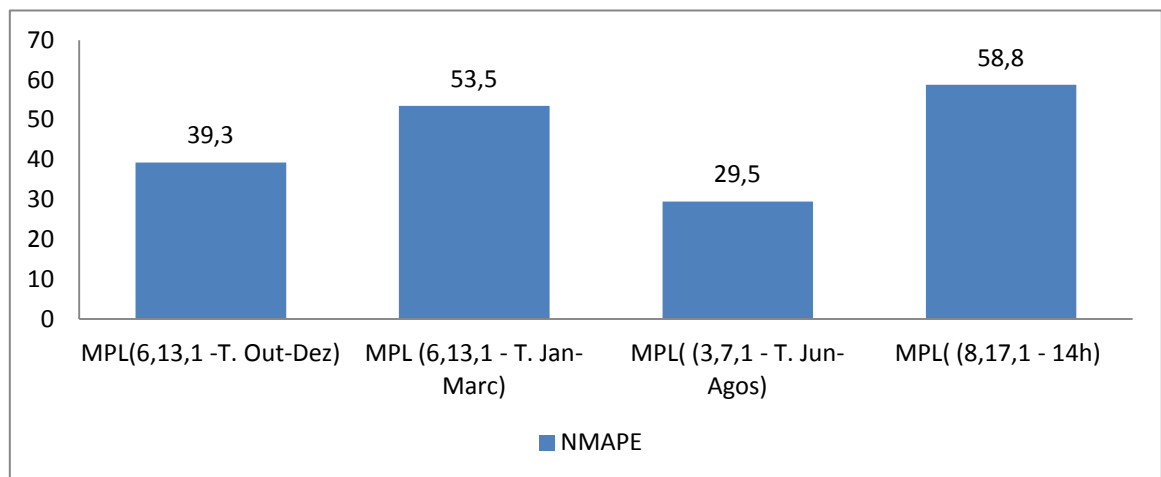


Fig. 29: Gráfico dos erros obtidos na previsão da potência pelas redes MLP.

Obviamente, o desempenho geral do modelo de previsão influencia o valor das medições dos erros. No entanto, o período também pode influenciar significativamente o desempenho aparente de um dado sistema de previsão assim como a variação dos parâmetros de entrada. O gráfico da figura 29 mostra os resultados obtidos com a variação dos parâmetros e entrada e a previsão para quatro períodos diferentes.

4.2.4.2. Função de Base Radial (RNAs-RBF)

Após a modelação da rede *MLP* e ter encontrado a rede com melhor desempenho *MLP* (3,7,1) foram utilizadas as variáveis e os intervalos de tempo que proporcionaram esses

resultados para a modelação da rede *RBF*. As variáveis foram: velocidade do vento, direção do vento e temperatura nos intervalos de tempo, janeiro a maio e de setembro a dezembro, conjunto de treino com 4639 (quatro mil seiscentos e trinta e nove observações), dos meses de junho a agosto, conjunto de teste, correspondente a 1813 (mil oitocentos e treze observações). Seguindo um procedimento comum nas aplicações de rede *RBF*, os valores dos conjuntos de treino e de teste foram normalizados antes do início do processo de treino da rede.

A fim de verificar o comportamento do modelo da *RNs-RBF* foram realizadas várias simulações e registadas, para não ser influenciado por uma inicialização específica e aleatória do peso. Os resultados estão apresentados na tabela que se segue:

Tabela 14: índices de desempenho da rede RBF.

| Input | Simulação | NME | NMAE | NMAPE | NRMSE | Correl. |
|-------|-----------|--------|-------|--------|--------|---------|
| VDVT | 1 | -0,183 | 0,387 | 38,743 | 0,4925 | 0,605 |
| VDVT | 2 | -0,203 | 0,404 | 40,452 | 0,5158 | 0,597 |
| VDVT | 3 | -0,182 | 0,385 | 38,594 | 0,490 | 0,599 |
| VDVT | 4 | -0,194 | 0,389 | 38,994 | 0,4952 | 0,601 |
| VDVT | 5 | -0,151 | 0,359 | 35,903 | 0,4510 | 0,612 |
| VDVT | 6 | -0,185 | 0,385 | 38,518 | 0,488 | 0,607 |

Para ter uma ideia gráfica dos erros obtidos pela rede, seguem as Fig. 30 e Fig. 31. Nenhum dos erros de previsão, erro médio normalizado, erro médio absoluto normalizado, erro quadrático médio normalizado e raiz quadrada do erro médio absoluto percentual normalizado convergem para o erro meta estipulado (0,01). Os modelos não capturaram adequadamente o processo gerador da série. Verifica-se isto, pois, os valores dos erros, absoluto médio percentual normalizado, estão entre 35,9 % e 40,5 %.

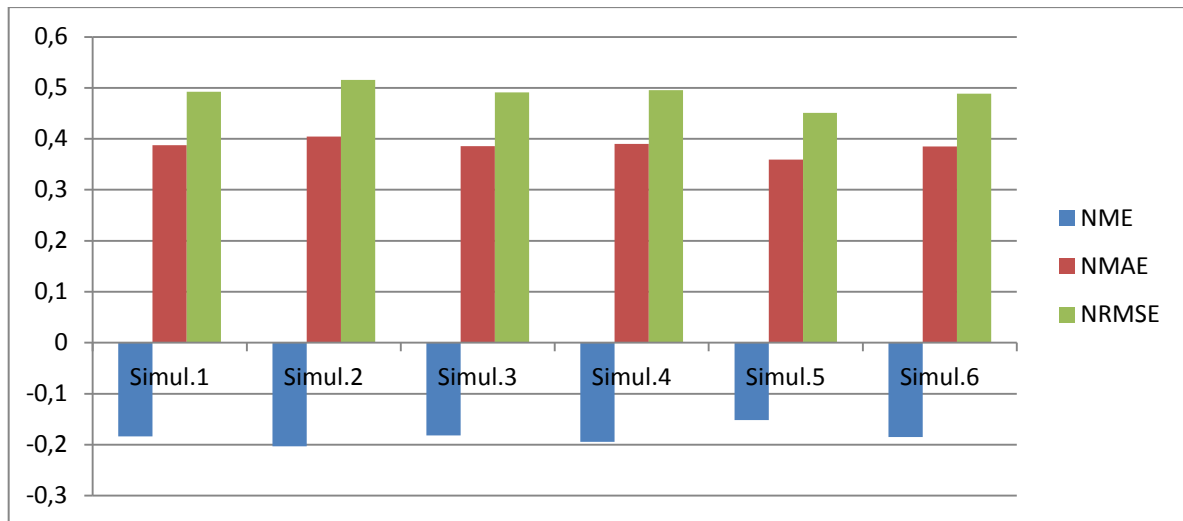


Fig. 30: Gráfico dos índices de desempenho da rede RBF.

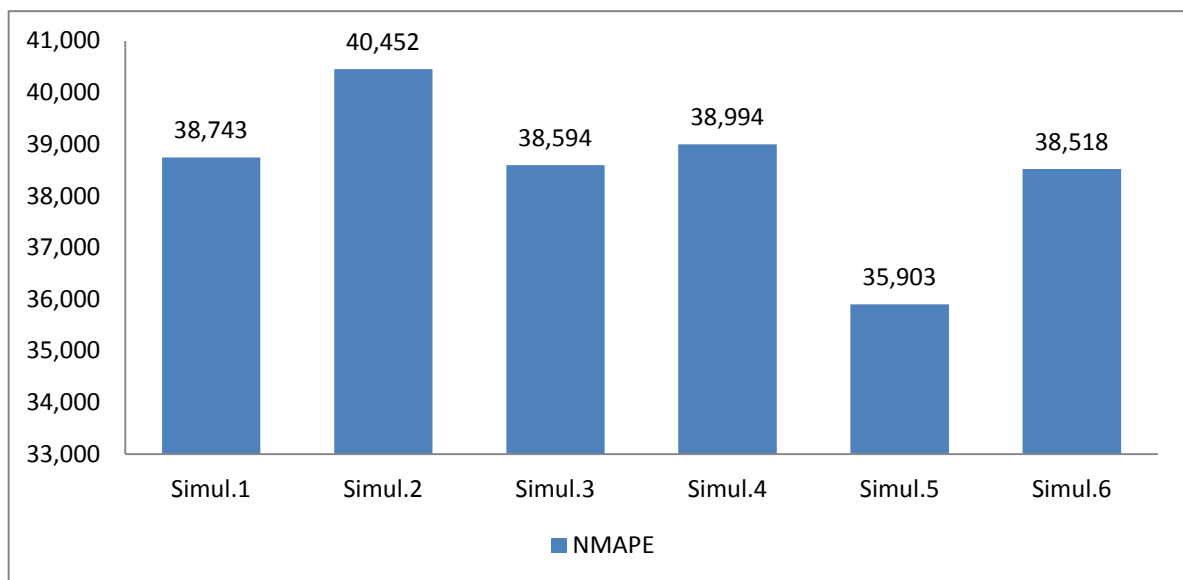


Fig. 31: Gráfico dos erros obtidos na previsão da potência pelas redes RBF nas diferentes simulações.

O gráfico da figura 32 mostra a soma do erro quadrado (SSE), isto é, a soma dos quadrados dos erros de todos os padrões para todas as épocas.

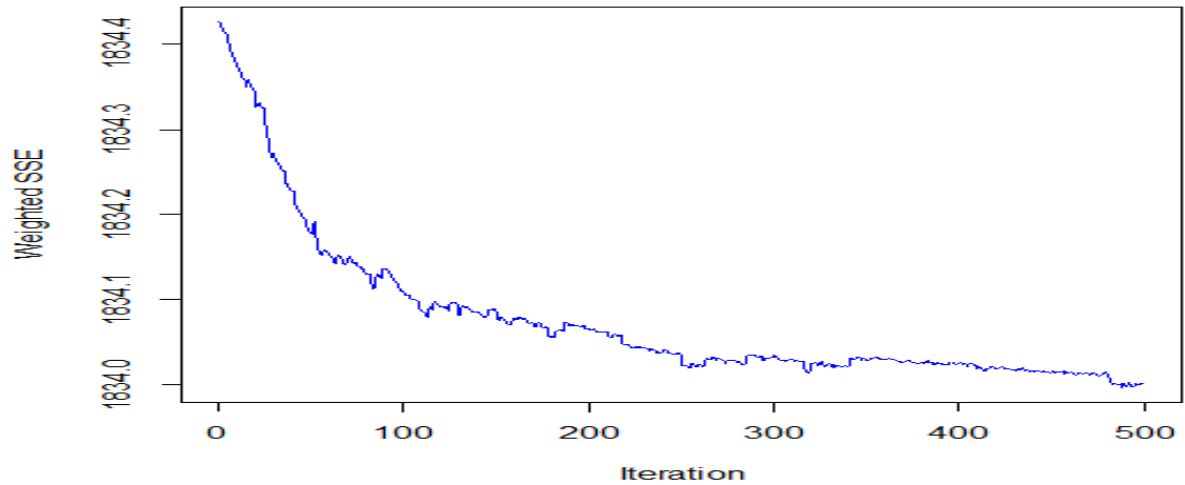


Fig. 32: Gráfico do treino da rede RBF.

Os gráficos que se seguem ilustram as previsões obtidas pelas redes *RBF*, de acordo com as simulações realizadas:

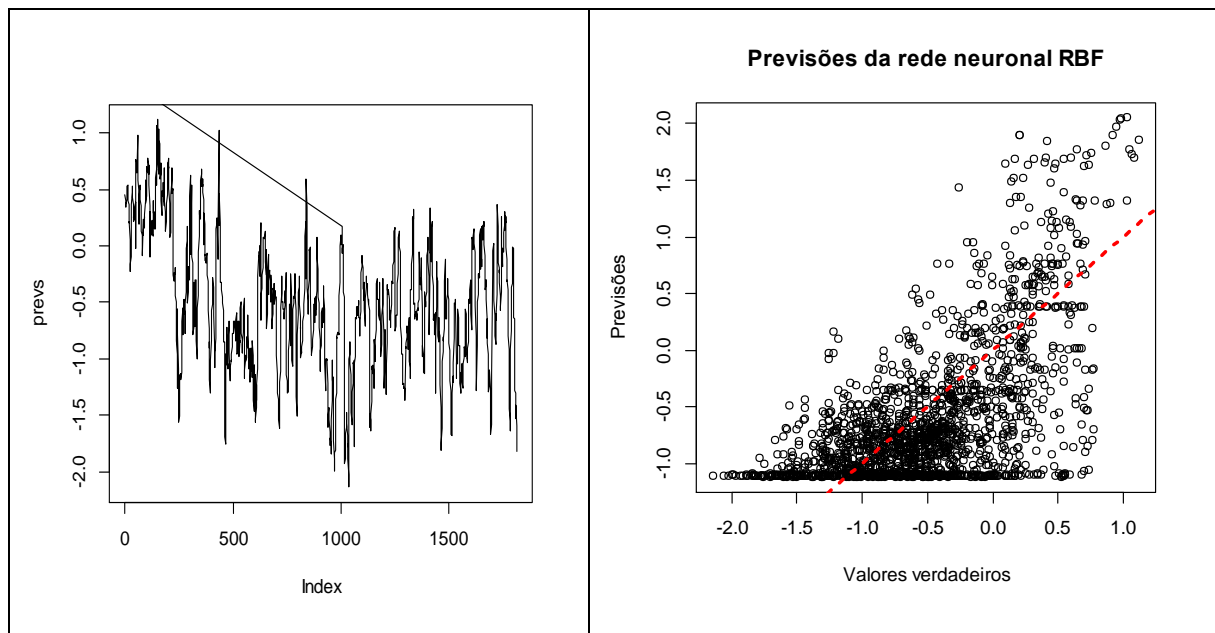


Fig. 33: Gráfico do comportamento do modelo RBF, simulação 1.

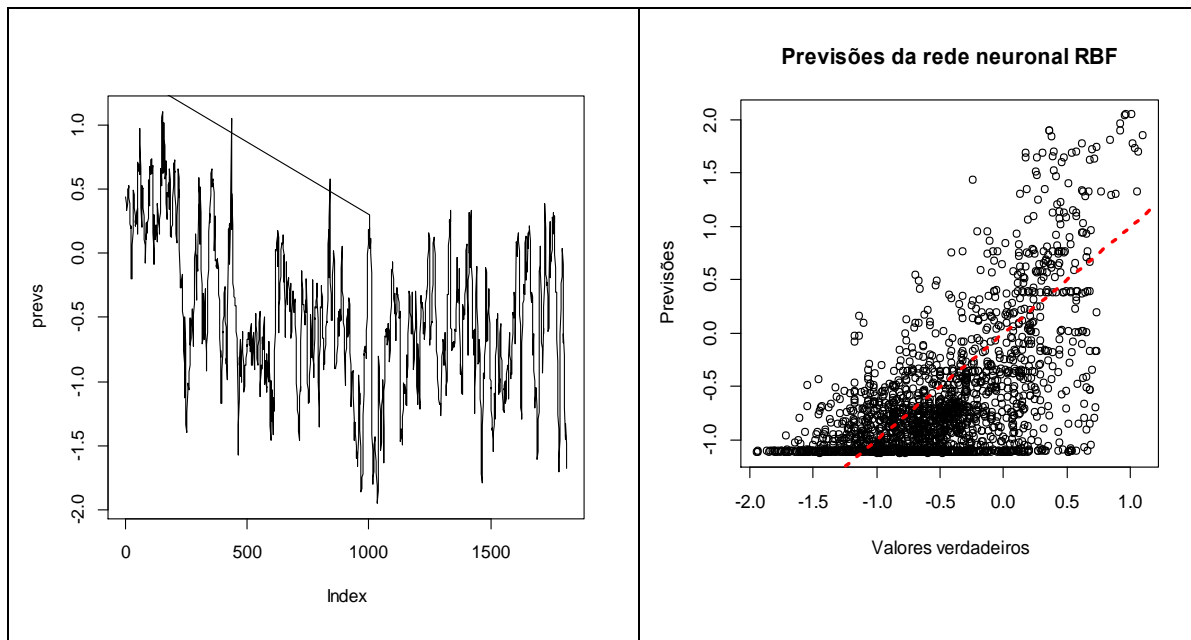


Fig. 34: Gráfico do comportamento do modelo RBF, simulação 2.

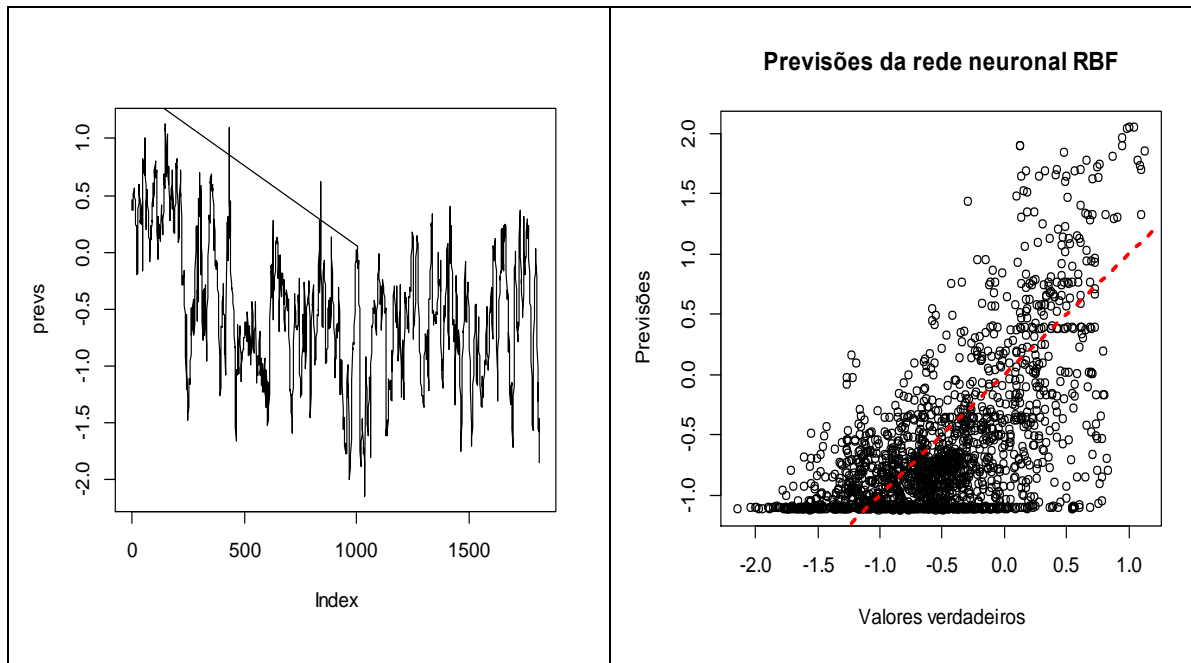


Fig. 35: Gráfico do comportamento do modelo RBF, simulação 3.

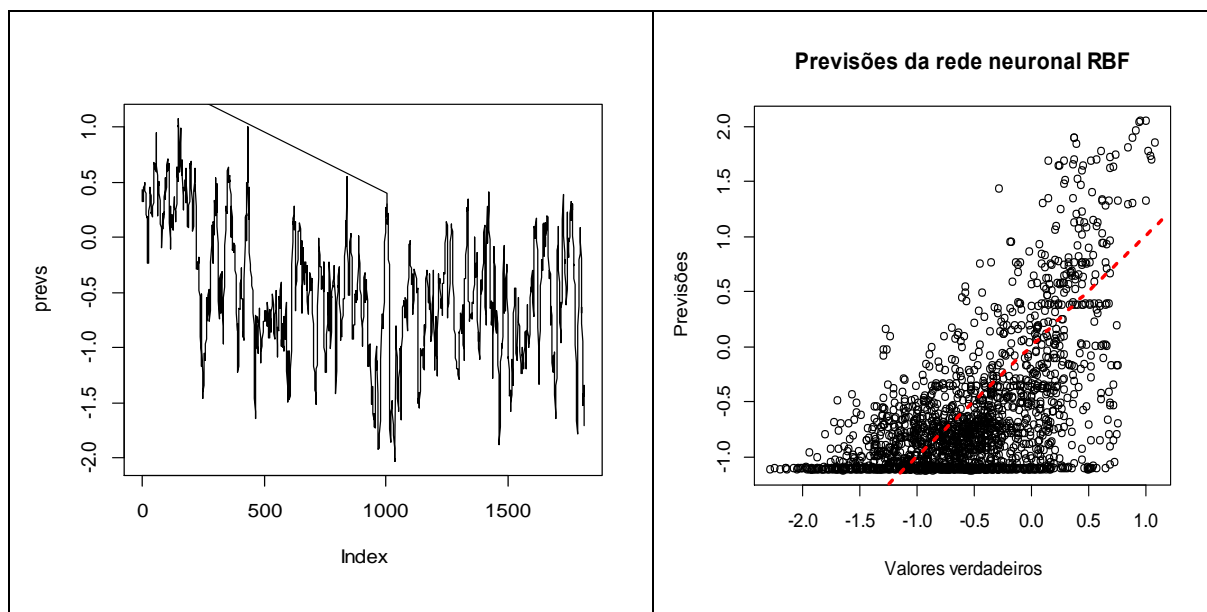


Fig. 36: Gráfico do comportamento do modelo RBF, simulação 4.

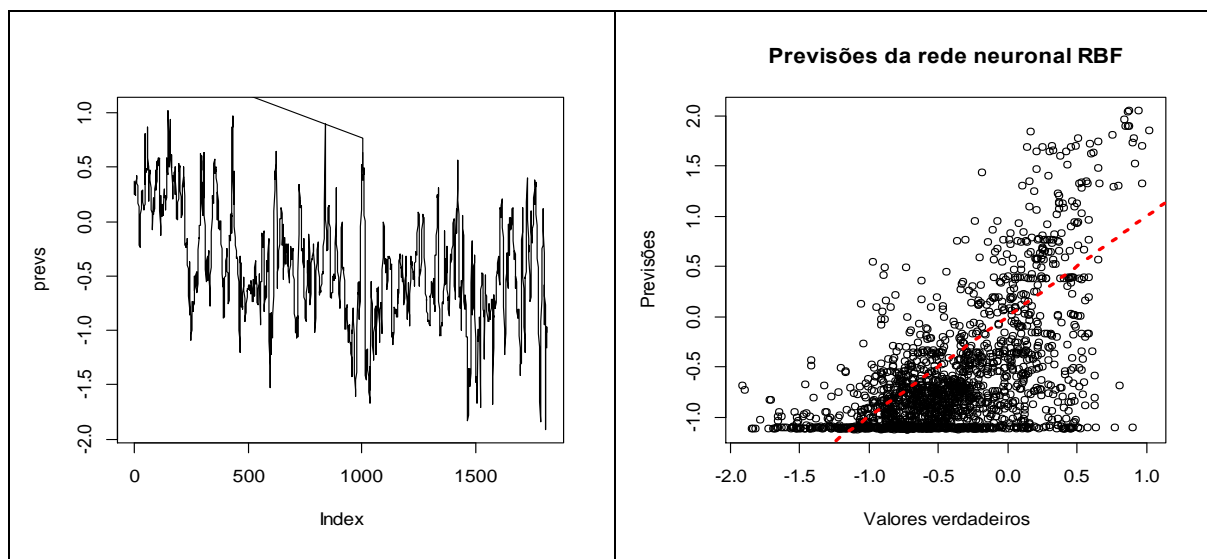


Fig. 37: Gráfico do comportamento do modelo RBF, simulação 5.

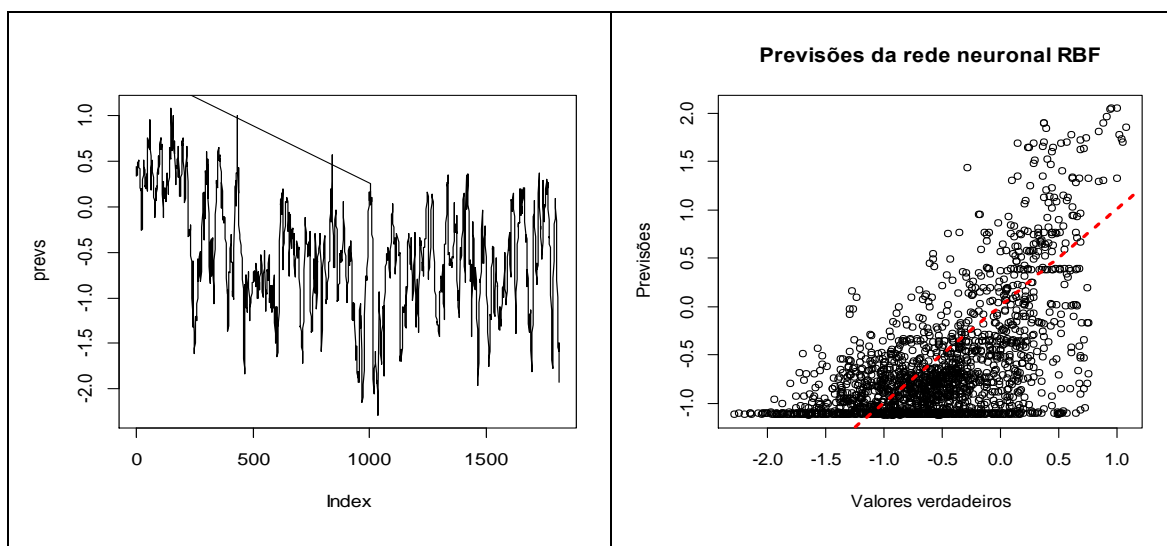


Fig. 38: Gráfico do comportamento do modelo RBF, simulação 6.

4.2.5. Discussão dos Resultados Obtidos

A discussão é feita com base nos resultados obtidos com a realização dos diferentes testes, ou seja primeiramente serão relatados os fatos verificados com a utilização do modelo *Holt-Winters*, em seguida os fatos verificados com a utilização dos modelos *ARIMA* e por último os fatos verificados com a utilização dos modelos de *RNs*.

A tabela que se segue apresenta os índices de desempenho de cada modelo estudado para as séries com os melhores desempenhos:

Tabela 15: Índice do desempenho de cada modelo utilizado na modelação das séries.

| Índice do desempenho | Modelos | | | |
|----------------------|--------------|--------|----------------------|----------------------|
| | Holt-Winters | ARIMA | RN _s -MLP | RN _s -RBF |
| ME | -0,301 | -0,004 | -0,175 | -0,151 |
| MAPE | 76,202 | 51,351 | 29,528 | 35,902 |

O modelo *Holt-Winters* não apresenta resultado satisfatório em termos de previsão. A magnitude do erro deste modelo foi superior aos demais utilizados.

Quanto ao modelo *ARIMA* verificou-se que este apresenta resultado mais satisfatório em relação ao modelo *Holt-Winters*. Porém em relação aos modelos de *RNs* este fato não é verdadeiro.

Observa-se que o modelo genérico *RNs* foi o que apresentou o melhor resultado em termos de capacidade preditiva, sendo que o modelo *MLP (3,7,1)* foi aquele que apresentou o melhor resultado dentro desta categoria.

5. Conclusões

As informações obtidas de um parque de produção de eletricidade de base eólico são importantes para prever, planejar e ajudar na tomada de decisões. A utilização dos métodos de previsão de séries temporais para prever a produção da energia elétrica a partir da energia eólica em Cabo Verde, particularmente na ilha de Santiago, utilizando as informações/dados do parque, permitiu analisar os resultados das previsões obtidas e comparar os resultados dos diferentes modelos utilizados nos estudos das séries temporais.

Todas as séries utilizadas tiveram o mesmo período, janeiro a dezembro de 2012, nove meses para treino e três para teste, com variações dos meses utilizados. No entanto essas variações fizeram com que os elementos de comparação dos índices de desempenho dos modelos sejam ME e MAPE.

O primeiro modelo utilizado, modelo de suavização exponencial de *Holt Winters* é explicado pelo fato da série apresentar tendência e sazonalidade. Nesta modelação priorizou-se o modelo multiplicativo cuja amplitude da variação sazonal varia ao longo do tempo. O segundo modelo, o *ARIMA* que possui apenas uma variável, a potência produzida pelo parque eólico descreve o comportamento dessa variável em termos de seus valores passados. O terceiro modelo, as redes neuronais artificiais por terem mostrado grandes desempenhos nas previsões de séries temporais.

Os modelos foram submetidos a uma busca pelas melhores variáveis de entrada, que proporcionaram menores erros de previsão, sendo cada uma das séries tratada separadamente.

O modelo de redes neuronais perceptron de multicamadas, *RNs-MLP* foi comparativamente com os outros modelos estudados, o que apresenta melhor desempenho.

Um aspeto a ter em consideração é o histórico de dados utilizados nos modelos, uma vez que, para funcionarem com um elevado nível de confiança é necessário apresentar um histórico de dados significativos. Pois o parque da ilha de Santiago tem relativamente pouco tempo de existência e pouco tempo de produção (inaugurado em outubro de 2011). Com isso quer se dizer o seguinte: considerando a integração da energia eólica numa escala

de médio a grande porte, uma novidade em Cabo Verde, então esses aproximadamente 2 anos de existência do parque de certa forma se tornou como base de aprendizado constante, e com isso o parque sempre trabalhou pode assim dizer muito condicionado visto o grande potencial eólico existente, por outras palavras sempre produziu abaixo do que realmente poderia estar produzindo.

Tem-se notado dois períodos de vento na ilha de Santiago em que em alguns meses os ventos são fortes e outros são fracos. Logo, o período também, definição do conjunto de teste nos meses de junho a agosto influenciou significativamente no desempenho do modelo de previsão, redes neuronais, assim como a variação dos parâmetros de entrada.

Uma vez que a empresa Cabeólica, empresa que comercializa a energia para a ELECTRA, empresa que por sua vez fornece energia elétrica aos utentes, ainda não usa modelos de previsão da potência da produção da energia elétrica com base eólica os modelos utilizados neste trabalho poderão servir de modelos de teste a quando da implementação desse serviço em Cabo Verde.

Almejando atingir um cenário de 50% de energias renováveis em 2020 e com principal enfoque na energia eólica, cenário bastante ambicioso, diga-se de passagem, se se tiver ainda em consideração a atual conjuntura de crise internacional, é necessário que se desenvolva alguns trabalhos futuros relacionados de forma a dar consistência ao trabalho ora apresentado. Para isso seguem algumas propostas que seriam interessantes considerar: utilizar mais de uma camada oculta na *RNs-MLP*, bem como outras funções de ativação do neurónio, e variação dos parâmetros, tais como o ajuste de peso aumentando o número de iterações; no modelo *RNs-RBF* o teste de outras funções radiais, como multi-quadrática, bem como outros métodos de treino e deteção de centros. Para as séries temporais também seria interessante estudar as mesmas séries agregadas aos dados obtidos nos anos subsequentes com frequências menores e maiores que 60 minutos.

A utilização de modelo híbrido, modelo que associa a elevada precisão dos modelos estatísticos, das séries temporais, na previsão para curto prazo com a precisão dos modelos físicos para previsões, como também modelos combinados de previsões (modelos

múltiplos) cujo princípio refere-se à possibilidade de aumentar a precisão das previsões, beneficiando-se da complementaridade da informação contida em cada previsão individual.

6. Referências Bibliográficas

- [1] EUROPEAN WIND ENERGY ASSOCIATION (EWEA); GREENPEACE. (2005). Wind force 12: a blueprint to achieve 12% of the world's electricity from windpower by 2020. London: EWEA.
- [2] INSTITUTO NACIONAL DE ESTATÍSTICA DE CABO VERDE. (2010). Relatório dos resultados preliminares do RGPH 2010. INE - Instituto Nacional de Estatística de Cabo Verde. Praia, pg. 1-26.
- [3] GONÇALVES FERNANDES, A., E. (1996). *Estudo de tensão para cenário de expansão do sistema elétrico da ilha de Santiago*. Trabalho Conclusão de Curso apresentado à disciplina Projeto de Graduação como requisito parcial à conclusão do Curso de Engenharia Eletrotécnica, setor de ciências tecnológicas, Engenharia Elétrica. Universidade Federal do Paraná, pg. 28.
- [4] ELECTRA. Relatório Anual Electra. Electra-Empresa de Eletricidade e Água, 2011.
- [5] ELECTRA. Relatório Anual Electra. Electra-Empresa de Eletricidade e Água, 2008.
- [6] Estudo de Impacte Ambiental, Cape Verde Wind Farm Extension Project, Fevereiro de 2009.
- [7] R.N.G.S.B.B. (2008). *Treino on Line de Redes Neurais com Critérios de Informação Aplicada à Previsão Eólica*. Porto: FEP.
- [8] ELVES, E.R.M.S. (2012). *Escalonamento de Geradores Térmicos com Integração de Produção Eólica: Uma Abordagem Baseada em Programação Inteira Mista*. Porto: FEUP.
- [9] LOBO, MIGUEL GARCÍA. (2010). *Métodos de predicción de la generación agregada de energía eólica*. Tesis Doctoral, UNIVERSIDAD CARLOS III DE MADRID.

- [10] GAMA, J. *et al.* (2012). *Extração de Conhecimento de Dados - Data Mining, Silabo.*
- [11] SOUSA, J. N. M. S. (2007). *Previsão da Produção Elétrica em Parques Eólicos.* Dissertação submetida para satisfação parcial dos requisitos do grau de Mestre em Engenharia Eletrotécnica e de Computadores, Porto: FEUP.
- [12] A. Joensen T. Nielsen, *et al.* (1999). *A new reference for Wind Power Forecasting.* Wind Energy, vol. 1, pg. 29-34.
- [13] WIND POWER FORECASTING: STATE-OF-THE-ART, (2009). ARGONNE.
- [14] Box, G. E. P. e Jenkins, G. M., (1976). Time Series Analysis. Forecasting and Control. Revised Edition. Oakland. California: Holden Day.
- [15] Makridakis, S.G., *et al.* (1997). Forecasting: Methods and Applications.
- [16] Pessanha, J. F. M. (s. d.). *Previsão da Velocidade do Vento com Sistema de Inferência Neuro-fuzzy Adaptativo.* UERJ – Universidade do Estado do Rio de Janeiro, Departamento de Estatística – Instituto de Matemática e Estatística. Rio de Janeiro – RJ.
- [17] Custódio, R. S. (2009). *Energia eólica para produção de energia elétrica.* Eletrobras, Rio de Janeiro.
- [18] ROB J. HYNDMAN e GEORGE, A. (s. d.). *Previsões: métodos e aplicações.*
- [19] CHATFIELD, C. (2004). *The Analysis of Time Séries - An Introduction.*
- [20] MAKRIDAKIS *et al.* (1998). *Forecasting: Methods and Applications.*

- [21] LANGE, M. (2005). *On the Uncertainty of Wind Power Predictions - Analysis of the Forecast Accuracy and Statistical Distribution of Errors. Journal of Solar Energy Engineering*. Vol. 127: pg. 177- 184.
- [22] Corrêa, R., *et al.* (s. d.). *Previsão de séries de tempo na presença de mudança estrutural: Redes neurais artificiais e modelos estruturais*.
- [23] BRAGA, *et al.* (2011). *Redes Neurais Artificiais: Teoria e aplicações*. 2ª ed. R.J.
- [24] FLORES, J. H. F. (2009). *Comparação de modelos MLP/RNA e modelos Box-jenkins em séries temporais não lineares*. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Porto Alegre.
- [25] KIRSTEN, Heitor André. (2009). *Comparação entre os modelos Holtwinters e Redes Neurais para previsão de séries temporais financeiras*. Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Mestre em Engenharia de Produção e Sistemas.
- [26] CASTRO, F. (2001). *Redes neurais artificiais*. Porto Alegre, R. S: Pontifícia Universidade Católica do Rio Grande do Sul.
- [27] SANTOS, A. A. P. (2005). *Previsão não-linear da taxa de câmbio real/dólar utilizando redes neurais e sistemas nebulosos*. Dissertação (Mestrado em Economia), Programa de Pós-Graduação em Economia, Universidade Federal de Santa Catarina, Florianópolis, SC.
- [28] Madsen, H., *et al.* (2004). *A Protocol for Standardizing the Performance Evaluation of Short-Term Wind Power Prediction Models*.

7. Anexo

Anexo I: Localização do parque eólico da ilha de Santiago [Cabeólica Wind Farm Project]

